Fixed Point Theorems and Applications

VITTORINO PATA

Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano

vittorino.pata@polimi.it

Contents

Preface 2

Notation 3

1. FIXED POINT THEOREMS 5
   - The Banach contraction principle 5
   - Sequences of maps and fixed points 11
   - Fixed points of non-expansive maps 13
   - The Riesz mean ergodic theorem 15
   - The Brouwer fixed point theorem 17
   - The Schauder-Tychonoff fixed point theorem 19
   - The Markov-Kakutani theorem 26
   - The Kakutani-Ky Fan theorem 28
   - Notes on Chapter 1 31

2. SOME APPLICATIONS OF FIXED POINT THEOREMS 33
   - The implicit function theorem 33
   - Ordinary differential equations in Banach spaces 38
   - Semilinear equations of evolution 43
   - An abstract elliptic problem 45
   - The invariant subspace problem 46
   - Measure preserving maps on compact Hausdorff spaces 48
   - Invariant means on semigroups 49
   - Haar measures 51
   - Game theory 56
   - Notes on Chapter 2 59

Bibliography 60

Preface

Fixed point theory is a fascinating subject, with an enormous number of applications in various fields of mathematics. Maybe due to this transversal character, I have always experienced some difficulties to find a book (unless expressly devoted to fixed points) treating the argument in a unitary fashion. In most cases, I noticed that fixed points pop up when they are needed. On the contrary, I believe that they should deserve a relevant place in any general textbook, and particularly, in a functional analysis textbook. This is mainly the reason that made me decide to write down these notes. I tried to collect most of the significant results of the field, and then to present various related applications.

This work consists of two chapters which, although rather self-contained, are conceived to be part of a future functional analysis book that I hope to complete in the years to come. Thus some general background is needed to approach the next pages. The reader is in fact supposed to be familiar with measure theory, Banach and Hilbert spaces, locally convex topological vector spaces and, in general, with linear functional analysis.

Even if one cannot possibly find here all the information and the details contained in a dedicated treatise, I hope that these notes could provide a quite satisfactory overview of the topic, at least from the point of view of analysis.

Vittorino Pata

Notation

Given a normed space X, $x \in X$ and $r > 0$ we denote

$$\begin{aligned} B_X(x,r) &= \{y \in X : \|y - x\| < r\} \\ \overline{B}_X(x,r) &= \{y \in X : \|y - x\| \leq r\} \\ \partial B_X(x,r) &= \{y \in X : \|y - x\| = r\}. \end{aligned}$$

Whenever misunderstandings might occur, we write $\|x\|_X$ to stress that the norm is taken in X. For a subset $Y \subset X$, we denote by $\overline{Y}$ the closure of Y, by Y^C the complement of Y, by $\text{span}(Y)$ the linear space generated by Y, and by $\text{co}(Y)$ the *convex hull* of Y, that is, the set of all finite convex combinations of elements of Y.

We will often use the notion of *uniformly convex* Banach space. Recall that a Banach space X is uniformly convex if given any two sequences $x_n, y_n \in X$ with

$$\|x_n\| \leq 1, \qquad \|y_n\| \leq 1, \qquad \lim_{n\to\infty} \|x_n + y_n\| = 2$$

it follows that

$$\lim_{n\to\infty} \|x_n - y_n\| = 0.$$

In particular, we will exploit the property (coming directly from the definition of uniform convexity) that minimizing sequences in closed convex subsets are convergent. Namely, if $C \subset X$ is (nonvoid) closed and convex and $x_n \in C$ is such that

$$\lim_{n\to\infty} \|x_n\| = \inf_{y\in C} \|y\|$$

then there exists a unique $x \in C$ such that

$$\|x\| = \inf_{y\in C} \|y\| \qquad \text{and} \qquad \lim_{n\to\infty} x_n = x.$$

Clearly, being Hilbert spaces uniformly convex, all the results involving uniformly convex Banach spaces can be read in terms of Hilbert spaces.

A weaker notion is *strict convexity*: a Banach space X is strictly convex if for all $x, y \in X$ with $x \neq y$ the relation

$$\|x\| = \|y\| \leq 1$$

implies

$$\|x + y\| < 2.$$

It is immediate to check from the definitions that uniform convexity implies strict convexity.

Other spaces widely used here are *locally convex spaces*. A locally convex space X is a vector space endowed with a family $\mathcal{P}$ of *separating seminorms*. Hence for every element $x \in X$ there is a seminorm $p \in \mathcal{P}$ such that $p(x) = 0$. Therefore $\mathcal{P}$ gives X the structure of (Hausdorff) topological vector space in which there is a local base whose members are covex. A local base $\mathcal{B}$ for the topology is given by finite intersections of sets of the form

$$\big\{x \in X : p(x) < \varepsilon\big\}$$

for some $p \in \mathcal{P}$ and some $\varepsilon > 0$. Note that, given $U \in \mathcal{B}$, there holds

$$U + U := \big\{x + y : x \in U,\ y \in U\big\} = 2U := \big\{2x : x \in U\big\}.$$

Good references for these arguments are, e.g., the books [1, 4, 14]. Concerning measure theory, we address the reader to [11, 13].

1. FIXED POINT THEOREMS

Fixed point theorems concern maps f of a set X into itself that, under certain conditions, admit a *fixed point*, that is, a point $x \in X$ such that $f(x) = x$. The knowledge of the existence of fixed points has relevant applications in many branches of analysis and topology. Let us show for instance the following simple but indicative example.

1.1 Example Suppose we are given a system of n equations in n unknowns of the form

$$g_j(x_1, \ldots, x_n) = 0, \qquad j = 1, \ldots, n$$

where the g_j are continuous real-valued functions of the real variables x_j. Let $h_j(x_1, \ldots, x_n) = g_j(x_1, \ldots, x_n) + x_j$, and for any point $x = (x_1, \ldots, x_n)$ define $h(x) = (h_1(x), \ldots, h_n(x))$. Assume now that h has a fixed point $\bar{x} \in \mathbb{R}^n$. Then it is easily seen that $\bar{x}$ is a solution to the system of equations.

Various application of fixed point theorems will be given in the next chapter.

The Banach contraction principle

1.2 Definition Let X be a metric space equipped with a distance d. A map $f : X \to X$ is said to be *Lipschitz continuous* if there is $\lambda \geq 0$ such that

$$d(f(x_1), f(x_2)) \leq \lambda d(x_1, x_2), \qquad \forall\, x_1, x_2 \in X.$$

The smallest λ for which the above inequality holds is the *Lipschitz constant* of f. If $\lambda \leq 1$ f is said to be *non-expansive*, if $\lambda < 1$ f is said to be a *contraction*.

1.3 Theorem [Banach] *Let f be a contraction on a complete metric space X. Then f has a unique fixed point $\bar{x} \in X$.*

PROOF Notice first that if $x_1, x_2 \in X$ are fixed points of f, then

$$d(x_1, x_2) = d(f(x_1), f(x_2)) \leq \lambda d(x_1, x_2)$$

which imply $x_1 = x_2$. Choose now any $x_0 \in X$, and define the iterate sequence $x_{n+1} = f(x_n)$. By induction on n,

$$d(x_{n+1}, x_n) \leq \lambda^n d(f(x_0), x_0).$$

If $n \in \mathbb{N}$ and $m \geq 1$,

$$\begin{aligned} d(x_{n+m}, x_n) &\leq d(x_{n+m}, x_{n+m-1}) + \cdots + d(x_{n+1}, x_n) \\ &\leq (\lambda^{n+m} + \cdots + \lambda^n) d(f(x_0), x_0) \qquad (1) \\ &\leq \frac{\lambda^n}{1-\lambda} d(f(x_0), x_0). \end{aligned}$$

Hence x_n is a Cauchy sequence, and admits a limit $\bar{x} \in X$, for X is complete. Since f is continuous, we have $f(\bar{x}) = \lim_n f(x_n) = \lim_n x_{n+1} = \bar{x}$. $\diamond$

Remark Notice that letting $m \to \infty$ in (1) we find the relation

$$d(x_n, \bar{x}) \leq \frac{\lambda^n}{1-\lambda} d(f(x_0), x_0)$$

which provides a control on the convergence rate of x_n to the fixed point $\bar{x}$.

The completeness of X plays here a crucial role. Indeed, contractions on incomplete metric spaces may fail to have fixed points.

Example Let $X = (0, 1]$ with the usual distance. Define $f : X \to X$ as $f(x) = x/2$.

1.4 Corollary *Let X be a complete metric space and Y be a topological space. Let $f : X \times Y \to X$ be a continuous function. Assume that f is a contraction on X uniformly in Y, that is,*

$$d(f(x_1, y), f(x_2, y)) \leq \lambda d(x_1, x_2), \qquad \forall\, x_1, x_2 \in X,\ \forall\, y \in Y$$

for some $\lambda < 1$. Then, for every fixed $y \in Y$, the map $x \mapsto f(x, y)$ has a unique fixed point $\varphi(y)$. Moreover, the function $y \mapsto \varphi(y)$ is continuous from Y to X.

Notice that if $f : X \times Y \to X$ is continuous on Y and is a contraction on X uniformly in Y, then f is in fact continuous on $X \times Y$.

PROOF In light of Theorem 1.3, we only have to prove the continuity of φ. For $y, y_0 \in Y$, we have

$$\begin{aligned} d(\varphi(y), \varphi(y_0)) &= d(f(\varphi(y), y), f(\varphi(y_0), y_0)) \\ &\leq d(f(\varphi(y), y), f(\varphi(y_0), y)) + d(f(\varphi(y_0), y), f(\varphi(y_0), y_0)) \\ &\leq \lambda d(\varphi(y), \varphi(y_0)) + d(f(\varphi(y_0), y), f(\varphi(y_0), y_0)) \end{aligned}$$

which implies

$$d(\varphi(y), \varphi(y_0)) \leq \frac{1}{1-\lambda} d(f(\varphi(y_0), y), f(\varphi(y_0), y_0)).$$

Since the above right-hand side goes to zero as $y \to y_0$, we have the desired continuity. $\diamond$

Remark If in addition Y is a metric space and f is Lipschitz continuous in Y, uniformly with respect to X, with Lipschitz constant $L \geq 0$, then the function $y \mapsto \varphi(y)$ is Lipschitz continuous with Lipschitz constant less than or equal to $L/(1-\lambda)$.

Theorem 1.3 gives a sufficient condition for f in order to have a unique fixed point.

Example Consider the map

$$g(x) = \begin{cases} 1/2 + 2x & x \in [0, 1/4] \\ 1/2 & x \in (1/4, 1] \end{cases}$$

mapping $[0, 1]$ onto itself. Then g is not even continuous, but it has a unique fixed point ($x = 1/2$).

The next corollary takes into account the above situation, and provides existence and uniqueness of a fixed point under more general conditions.

Definition For $f : X \to X$ and $n \in \mathbb{N}$, we denote by f^n the n^{th}*-iterate* of f, namely, $f \circ \cdots \circ f$ n-times (f^0 is the identity map).

1.5 Corollary *Let X be a complete metric space and let $f : X \to X$. If f^n is a contraction, for some $n \geq 1$, then f has a unique fixed point $\bar{x} \in X$.*

PROOF Let $\bar{x}$ be the unique fixed point of f^n, given by Theorem 1.3. Then $f^n(f(\bar{x})) = f(f^n(\bar{x})) = f(\bar{x})$, which implies $f(\bar{x}) = \bar{x}$. Since a fixed point of f is clearly a fixed point of f^n, we have uniqueness as well. $\diamond$

Notice that in the example $g^2(x) \equiv 1/2$.

1.6 Further extensions of the contraction principle There is in the literature a great number of generalizations of Theorem 1.3 (see, e.g., [6]). Here we point out some results.

Theorem [Boyd-Wong] *Let X be a complete metric space, and let $f : X \to X$. Assume there exists a right-continuous function $\varphi : [0, \infty) \to [0, \infty)$ such that $\varphi(r) < r$ if $r > 0$, and*

$$d(f(x_1), f(x_2)) \leq \varphi(d(x_1, x_2)), \qquad \forall\, x_1, x_2 \in X.$$

Then f has a unique fixed point $\bar{x} \in X$. Moreover, for any $x_0 \in X$ the sequence $f^n(x_0)$ converges to $\bar{x}$.

Clearly, Theorem 1.3 is a particular case of this result, for $\varphi(r) = \lambda r$.

PROOF If $x_1, x_2 \in X$ are fixed points of f, then

$$d(x_1, x_2) = d(f(x_1), f(x_2)) \leq \varphi(d(x_1, x_2))$$

so $x_1 = x_2$. To prove the existence, fix any $x_0 \in X$, and define the iterate sequence $x_{n+1} = f(x_n)$. We show that x_n is a Cauchy sequence, and the desired conclusion follows arguing like in the proof of Theorem 1.3. For $n \geq 1$, define the positive sequence

$$a_n = d(x_n, x_{n-1}).$$

It is clear that $a_{n+1} \leq \varphi(a_n) \leq a_n$; therefore a_n converges monotonically to some $a \geq 0$. From the right-continuity of φ, we get $a \leq \varphi(a)$, which entails $a = 0$. If x_n is not a Cauchy sequence, there is $\varepsilon > 0$ and integers $m_k > n_k \geq k$ for every $k \geq 1$ such that

$$d_k := d(x_{m_k}, x_{n_k}) \geq \varepsilon, \qquad \forall\, k \geq 1.$$

In addition, upon choosing the smallest possible m_k, we may assume that

$$d(x_{m_{k-1}}, x_{n_k}) < \varepsilon$$

for k big enough (here we use the fact that $a_n \to 0$). Therefore, for k big enough,

$$\varepsilon \leq d_k \leq d(x_{m_k}, x_{m_{k-1}}) + d(x_{m_{k-1}}, x_{n_k}) < a_{m_k} + \varepsilon$$

implying that $d_k \to \varepsilon$ from above as $k \to \infty$. Moreover,

$$d_k \leq d_{k+1} + a_{m_{k+1}} + a_{n_{k+1}} \leq \varphi(d_k) + a_{m_{k+1}} + a_{n_{k+1}}$$

and taking the limit as $k \to \infty$ we obtain the relation $\varepsilon \leq \varphi(\varepsilon)$, which has to be false since $\varepsilon > 0$. ◇

Theorem [Caristi] *Let X be a complete metric space, and let $f : X \to X$. Assume there exists a lower semicontinuous function $\psi : X \to [0, \infty)$ such that*

$$d(x, f(x)) \leq \psi(x) - \psi(f(x)), \qquad \forall\, x \in X.$$

Then f has (at least) a fixed point in X.

Again, Theorem 1.3 is a particular case, obtained for $\psi(x) = d(x, f(x))/(1 - \lambda)$. Notice that f need not be continuous.

PROOF We introduce a partial ordering on X, setting

$$x \preceq y \qquad \text{if and only if} \qquad d(x, y) \leq \psi(x) - \psi(y).$$

Let $\emptyset \neq X_0 \subset X$ be totally ordered, and consider a sequence $x_n \in X_0$ such that $\psi(x_n)$ is decreasing to $\alpha := \inf\{\psi(x) : x \in X_0\}$. If $n \in \mathbb{N}$ and $m \geq 1$,

$$\begin{aligned} d(x_{n+m}, x_n) &\leq \sum_{i=0}^{m-1} d(x_{n+i+1}, x_{n+i}) \\ &\leq \sum_{i=0}^{m-1} \psi(x_{n+i}) - \psi(x_{n+i+1}) \\ &= \psi(x_n) - \psi(x_{n+m}). \end{aligned}$$

Hence x_n is a Cauchy sequence, and admits a limit $x_* \in X$, for X is complete. Since ψ can only jump downwards (being lower semicontinuous), we also have $\psi(x_*) = \alpha$. If $x \in X_0$ and $d(x, x_*) > 0$, then it must be $x \preceq x_n$ for large n. Indeed, $\lim_n \psi(x_n) = \psi(x_*) \leq \psi(x)$. We conclude that x_* is an upper bound for X_0, and by the Zorn lemma there exists a maximal element $\bar{x}$. On the other hand, $\bar{x} \preceq f(\bar{x})$, thus the maximality of $\bar{x}$ forces the equality $\bar{x} = f(\bar{x})$. $\diamond$

If we assume the continuity of f, we obtain a slightly stronger result, even relaxing the continuity hypothesis on ψ.

Theorem *Let X be a complete metric space, and let $f : X \to X$ be a continuous map. Assume there exists a function $\psi : X \to [0, \infty)$ such that*

$$d(x, f(x)) \leq \psi(x) - \psi(f(x)), \qquad \forall\, x \in X.$$

Then f has a fixed point in X. Moreover, for any $x_0 \in X$ the sequence $f^n(x_0)$ converges to a fixed point of f.

PROOF Choose $x_0 \in X$. Due the above condition, the sequence $\psi(f^n(x_0))$ is decreasing, and thus convergent. Reasoning as in the proof of the Caristi theorem, we get that $f^n(x_0)$ admits a limit $\bar{x} \in X$, for X is complete. The continuity of f then entails $f(\bar{x}) = \lim_n f(f^n(x_0)) = \bar{x}$. $\diamond$

We conclude with the following extension of Theorem 1.3, that we state without proof.

Theorem [Čirič] *Let X be a complete metric space, and let $f : X \to X$ be such that*

$$\begin{aligned} &d(f(x_1), f(x_2)) \\ &\quad \leq \lambda \max \Big\{ d(x_1, x_2), d(x_1, f(x_1)), d(x_2, f(x_2)), d(x_1, f(x_2)), d(x_2, f(x_1)) \Big\} \end{aligned}$$

for some $\lambda < 1$ and every $x_1, x_2 \in X$. Then f has a unique fixed point $\bar{x} \in X$. Moreover, $d(f^n(x_0), \bar{x}) = O(\lambda^n)$ for any $x_0 \in X$.

Also in this case f need not be continuous. However, it is easy to check that it is continuous at the fixed point. The function g of the former example fulfills the hypotheses of the theorem with $\lambda = 2/3$.

1.7 Weak contractions We now dwell on the case of maps on a metric space which are contractive without being contractions.

Definition Let X be a metric space with a distance d. A map $f : X \to X$ is a *weak contraction* if

$$d(f(x_1), f(x_2)) < d(x_1, x_2), \qquad \forall\, x_1 \neq x_2 \in X.$$

Being a weak contraction is not in general a sufficient condition for f in order to have a fixed point, as it is shown in the following simple example.

Example Consider the complete metric space $X = [1, +\infty)$, and let $f : X \to X$ be defined as $f(x) = x + 1/x$. It is easy to see that f is a weak contraction with no fixed points.

Nonetheless, the condition turns out to be sufficient when X is compact.

Theorem *Let f be a weak contraction on a compact metric space X. Then f has a unique fixed point $\bar{x} \in X$. Moreover, for any $x_0 \in X$ the sequence $f^n(x_0)$ converges to $\bar{x}$.*

PROOF The uniqueness argument goes exactly as in the proof of Theorem 1.3. From the compactness of X, the continuous function $x \mapsto d(x, f(x))$ attains its minimum at some $\bar{x} \in X$. If $\bar{x} \neq f(\bar{x})$, we get

$$d(\bar{x}, f(\bar{x})) = \min_{x \in X} d(x, f(x)) \leq d(f(\bar{x}), f(f(\bar{x}))) < d(\bar{x}, f(\bar{x}))$$

which is impossible. Thus $\bar{x}$ is the unique fixed point of f (and so of f^n for all $n \geq 2$). Let now $x_0 \neq \bar{x}$ be given, and define $d_n = d(f^n(x_0), \bar{x})$. Observe that

$$d_{n+1} = d(f^{n+1}(x_0), f(\bar{x})) < d(f^n(x_0), \bar{x}) = d_n.$$

Hence d_n is strictly decreasing, and admits a limit $r \geq 0$. Let now $f^{n_k}(x_0)$ be a subsequence of $f^n(x_0)$ converging to some $z \in X$. Then

$$r = d(z, \bar{x}) = \lim_{k \to \infty} d_{n_k} = \lim_{k \to \infty} d_{n_k+1} = \lim_{k \to \infty} d(f(f^{n_k}(x_0)), \bar{x}) = d(f(z), \bar{x}).$$

But if $z \neq \bar{x}$, then

$$d(f(z), \bar{x}) = d(f(z), f(\bar{x})) < d(z, \bar{x}).$$

Therefore any convergent subsequence of $f^n(x_0)$ has limit $\bar{x}$, which, along with the compactness of X, implies that $f^n(x_0)$ converges to $\bar{x}$. $\diamond$

Obviously, we can relax the compactness of X by requiring that $\overline{f(X)}$ be compact (just applying the theorem on the restriction of f on $\overline{f(X)}$).

Arguing like in Corollary 1.5, it is also immediate to prove the following

Corollary *Let X be a compact metric space and let $f : X \to X$. If f^n is a weak contraction, for some $n \geq 1$, then f has a unique fixed point $\bar{x} \in X$.*

1.8 A converse to the contraction principle Assume we are given a set X and a map $f : X \to X$. We are interested to find a metric d on X such that (X, d) is a complete metric space and f is a contraction on X. Clearly, in light of Theorem 1.3, a necessary condition is that each iterate f^n has a unique fixed point. Surprisingly enough, the condition turns out to be sufficient as well.

Theorem [Bessaga] *Let X be an arbitrary set, and let $f : X \to X$ be a map such that f^n has a unique fixed point $\bar{x} \in X$ for every $n \geq 1$. Then for every $\varepsilon \in (0,1)$, there is a metric $d = d_\varepsilon$ on X that makes X a complete metric space, and f is a contraction on X with Lipschitz constant equal to ε.*

PROOF Choose $\varepsilon \in (0,1)$. Let Z be the subset of X consisting of all elements z such that $f^n(z) = \bar{x}$ for some $n \in \mathbb{N}$. We define the following equivalence relation on $X \setminus Z$: we say that $x \sim y$ if and only if $f^n(x) = f^m(y)$ for some $n, m \in \mathbb{N}$. Notice that if $f^n(x) = f^m(y)$ and $f^{n'}(x) = f^{m'}(y)$ then $f^{n+m'}(x) = f^{m+n'}(x)$. But since $x \notin Z$, this yields $n + m' = m + n'$, that is, $n - m = n' - m'$. At this point, by means of the axiom of choice, we select an element from each equivalence class. We now proceed defining the distance of $\bar{x}$ from a generic $x \in X$ by setting $d(\bar{x}, \bar{x}) = 0$, $d(x, \bar{x}) = \varepsilon^{-n}$ if $x \in Z$ with $x \neq \bar{x}$, where $n = \min\{m \in \mathbb{N} : f^m(x) = \bar{x}\}$, and $d(x, \bar{x}) = \varepsilon^{n-m}$ if $x \notin Z$, where $n, m \in \mathbb{N}$ are such that $f^n(\hat{x}) = f^m(x)$, $\hat{x}$ being the selected representative of the equivalence class $[x]$. The definition is unambiguous, due to the above discussion. Finally, for any $x, y \in X$, we set

$$d(x,y) = \begin{cases} d(x,\bar{x}) + d(y,\bar{x}) & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases}$$

It is straightforward to verify that d is a metric. To see that d is complete, observe that the only Cauchy sequences which do not converge to $\bar{x}$ are ultimately constant. We are left to show that f is a contraction with Lipschitz constant equal to ε. Let $x \in X$, $x \neq \bar{x}$. If $x \in Z$ we have

$$d(f(x), f(\bar{x})) = d(f(x), \bar{x}) \leq \varepsilon^{-n} = \varepsilon\varepsilon^{-(n+1)} = \varepsilon d(x, \bar{x}).$$

If $x \notin Z$ we have

$$d(f(x), f(\bar{x})) = d(f(x), \bar{x}) = \varepsilon^{n-m} = \varepsilon\varepsilon^{n-(m+1)} = \varepsilon d(x, \bar{x})$$

since $x \sim f(x)$. The thesis follows directly from the definition of the distance. $\diamond$

Sequences of maps and fixed points

Let (X, d) be a complete metric space. We consider the problem of convergence of fixed points for a sequence of maps $f_n : X \to X$. Corollary 1.5 will be implicitly used in the statements of the next two theorems.

1.9 Theorem *Assume that each f_n has at least a fixed point $x_n = f_n(x_n)$. Let $f : X \to X$ be a uniformly continuous map such that f^m is a contraction for some $m \geq 1$. If f_n converges uniformly to f, then x_n converges to $\bar{x} = f(\bar{x})$.*

PROOF We first assume that f is a contraction (i.e., $m = 1$). Let $\lambda < 1$ be the Lipschitz constant of f. Given $\varepsilon > 0$, choose $n_0 = n_0(\varepsilon)$ such that

$$d(f_n(x), f(x)) \leq \varepsilon(1 - \lambda), \qquad \forall\, n \geq n_0, \forall\, x \in X.$$

Then, for $n \geq n_0$,

$$\begin{aligned} d(x_n, \bar{x}) &= d(f_n(x_n), f(\bar{x})) \\ &\leq d(f_n(x_n), f(x_n)) + d(f(x_n), f(\bar{x})) \\ &\leq \varepsilon(1-\lambda) + \lambda d(x_n, \bar{x}). \end{aligned}$$

Therefore $d(x_n, \bar{x}) \leq \varepsilon$, which proves the convergence.
To prove the general case it is enough to observe that if

$$d(f^m(x), f^m(y)) \leq \lambda^m d(x, y)$$

for some $\lambda < 1$, we can define a new metric d_0 on X equivalent to d by setting

$$d_0(x, y) = \sum_{k=0}^{m-1} \frac{1}{\lambda^k} d(f^k(x), f^k(y)).$$

Moreover, since f is uniformly continuous, f_n converges uniformly to f also with respect to d_0. Finally, f is a contraction with respect to d_0. Indeed,

$$\begin{aligned} d_0(f(x), f(y)) &= \sum_{k=0}^{m-1} \frac{1}{\lambda^k} d(f^{k+1}(x), f^{k+1}(y)) \\ &= \lambda \sum_{k=1}^{m-1} \frac{1}{\lambda^k} d(f^k(x), f^k(y)) + \frac{1}{\lambda^{m+1}} d(f^m(x), f^m(y)) \\ &\leq \lambda \sum_{k=0}^{m-1} \frac{1}{\lambda^k} d(f^k(x), f^k(y)) = \lambda d_0(x, y). \end{aligned}$$

So the problem is reduced to the previous case $m = 1$. ◇

The next result refers to a special class of complete metric spaces.

1.10 Theorem *Let X be locally compact. Assume that for each $n \in \mathbb{N}$ there is $m_n \geq 1$ such that $f_n^{m_n}$ is a contraction. Let $f : X \to X$ be a map such that f^m is a contraction for some $m \geq 1$. If f_n converges pointwise to f, and f_n is an equicontinuous family, then $x_n = f_n(x_n)$ converges to $\bar{x} = f(\bar{x})$.*

PROOF Let $\varepsilon > 0$ be sufficiently small such that

$$K(\bar{x}, \varepsilon) := \{x \in X : d(x, \bar{x}) \leq \varepsilon\} \subset X$$

is compact. As a byproduct of the Ascoli theorem, f_n converges to f uniformly on $K(\bar{x}, \varepsilon)$, since it is equicontinuous and pointwise convergent. Choose $n_0 = n_0(\varepsilon)$ such that

$$d(f_n^m(x), f^m(x)) \leq \varepsilon(1-\lambda), \qquad \forall\, n \geq n_0, \forall\, x \in K(\bar{x}, \varepsilon)$$

where $\lambda < 1$ is the Lipschitz constant of f^m. Then, for $n \geq n_0$ and $x \in K(\bar{x}, \varepsilon)$ we have

$$\begin{aligned} d(f_n^m(x), \bar{x}) &= d(f_n^m(x), f^m(\bar{x})) \\ &\leq d(f_n^m(x), f^m(x)) + d(f^m(x), f^m(\bar{x})) \\ &\leq \varepsilon(1-\lambda) + \lambda d(x, \bar{x}) \leq \varepsilon. \end{aligned}$$

Hence $f_n^m(K(\bar{x}, \varepsilon)) \subset K(\bar{x}, \varepsilon)$ for all $n \geq n_0$. Since the maps $f_n^{m_n}$ are contractions, it follows that, for $n \geq n_0$, the fixed points x_n of f_n belong to $K(\bar{x}, \varepsilon)$, that is, $d(x_n, \bar{x}) \leq \varepsilon$. $\diamond$

Fixed points of non-expansive maps

Let X be a Banach space, $C \subset X$ nonvoid, closed, bounded and convex, and let $f : C \to C$ be a non-expansive map. The problem is whether f admits a fixed point in C. The answer, in general, is false.

Example Let $X = c_0$ with the supremum norm. Setting $C = \overline{B}_X(0,1)$, the map $f : C \to C$ defined by

$$f(x) = (1, x_0, x_1, \ldots), \qquad \text{for } x = (x_0, x_1, x_2, \ldots) \in C$$

is non-expansive but clearly admits no fixed points in C.

Things are quite different in uniformly convex Banach spaces.

1.11 Theorem [Browder-Kirk] *Let X be a uniformly convex Banach space and $C \subset X$ be nonvoid, closed, bounded and convex. If $f : C \to C$ is a non-expansive map, then f has a fixed point in C.*

We provide the proof in the particular case when X is a Hilbert space (the general case may be found, e.g., in [6], Ch.6.4).

PROOF Let $x_* \in C$ be fixed, and consider a sequence $r_n \in (0,1)$ converging to 1. For each $n \in \mathbb{N}$, define the map $f_n : C \to C$ as

$$f_n(x) = r_n f(x) + (1 - r_n) x_*.$$

Notice that f_n is a contractions on C, hence there is a unique $x_n \in C$ such that $f_n(x_n) = x_n$. Since C is weakly compact, x_n has a subsequence (still denoted by x_n) weakly convergent to some $\bar{x} \in C$. We shall prove that $\bar{x}$ is a fixed point of f. Notice first that

$$\lim_{n\to\infty} \left(\|f(\bar{x}) - x_n\|^2 - \|\bar{x} - x_n\|^2 \right) = \|f(\bar{x}) - \bar{x}\|^2.$$

Since f is non-expansive we have

$$\begin{aligned}\|f(\bar{x}) - x_n\| &\leq \|f(\bar{x}) - f(x_n)\| + \|f(x_n) - x_n\| \\ &\leq \|\bar{x} - x_n\| + \|f(x_n) - x_n\| \\ &= \|\bar{x} - x_n\| + (1 - r_n)\|f(x_n) - x_*\|.\end{aligned}$$

But $r_n \to 1$ as $n \to \infty$ and C is bounded, so we conclude that

$$\limsup_{n\to\infty} \left(\|f(\bar{x}) - x_n\|^2 - \|\bar{x} - x_n\|^2\right) \leq 0$$

which yields the equality $f(\bar{x}) = \bar{x}$. ◇

Proposition *In the hypotheses of Theorem 1.11, the set F of fixed points of f is closed and convex.*

PROOF The first assertion is trivial. Assume then $x_0, x_1 \in F$, with $x_0 \neq x_1$, and denote $x_t = (1 - t)x_0 + tx_1$, with $t \in (0, 1)$. We have

$$\|f(x_t) - x_0\| = \|f(x_t) - f(x_0)\| \leq \|x_t - x_0\| = t\|x_1 - x_0\|$$

$$\|f(x_t) - x_1\| = \|f(x_t) - f(x_1)\| \leq \|x_t - x_1\| = (1 - t)\|x_1 - x_0\|$$

that imply the equalities

$$\begin{aligned}\|f(x_t) - x_0\| &= t\|x_1 - x_0\| \\ \|f(x_t) - x_1\| &= (1 - t)\|x_1 - x_0\|.\end{aligned}$$

The proof is completed if we show that $f(x_t) = (1-t)x_0+tx_1$. This follows from a general fact about uniform convexity, which is recalled in the next lemma. ◇

Lemma *Let X be a uniformly convex Banach space, and let $\alpha, x, y \in X$ be such that*

$$\|\alpha - x\| = t\|x - y\|, \qquad \|\alpha - y\| = (1 - t)\|x - y\|,$$

for some $t \in [0, 1]$. Then $\alpha = (1 - t)x + ty$.

PROOF Without loss of generality, we can assume $t \geq 1/2$. We have

$$\begin{aligned}\|(1 - t)(\alpha - x) - t(\alpha - y)\| &= \|(1 - 2t)(\alpha - x) - t(x - y)\| \\ &\geq t\|x - y\| - (1 - 2t)\|\alpha - x\| \\ &= 2t(1 - t)\|x - y\|.\end{aligned}$$

Since the reverse inequality holds as well, and

$$(1 - t)\|\alpha - x\| = t\|\alpha - y\| = t(1 - t)\|x - y\|$$

from the uniform convexity of X (but strict convexity would suffice) we get

$$\|\alpha - (1 - t)x - ty\| = \|(1 - t)(\alpha - x) + t(\alpha - y)\| = 0$$

as claimed. ◇

The Riesz mean ergodic theorem

If T is a non-expansive linear map of a uniformly convex Banach space, then *all* the fixed points of T are recovered by means of a limit procedure.

1.12 Projections Let X be a linear space. A linear operator $P : X \to X$ is called a *projection* in X if $P^2x = PPx = Px$ for every $x \in X$. It is easy to check that P is the identity operator on $\operatorname{ran}(P)$, and the relations $\ker(P) = \operatorname{ran}(\mathbb{I} - P)$, $\operatorname{ran}(P) = \ker(\mathbb{I} - P)$ and $\ker(P) \cap \operatorname{ran}(P) = \{0\}$ hold. Moreover every element $x \in X$ admits a unique decomposition $x = y + z$ with $y \in \ker(P)$ and $z \in \operatorname{ran}(P)$.

> **Proposition** *If X is a Banach space, then a projection P is continuous if and only if $X = \ker(P) \oplus \operatorname{ran}(P)$.*

The notation $X = A \oplus B$ is used to mean that A and B are closed subspaces of X such that $A \cap B = \{0\}$ and $A + B = X$.

> PROOF If P is continuous, so is $\mathbb{I} - P$. Hence $\ker(P)$ and $\operatorname{ran}(P) = \ker(\mathbb{I} - P)$ are closed. Conversely, let $x_n \to x$, and $Px_n \to y$. Since $\operatorname{ran}(P)$ is closed, $y \in \operatorname{ran}(P)$, and therefore $Py = y$. But $Px_n - x_n \in \ker(P)$, and $\ker(P)$ is closed. So we have $x - y \in \ker(P)$, which implies $Py = Px$. From the closed graph theorem, P is continuous. ◇

1.13 Theorem [F. Riesz] *Let X be a uniformly convex Banach space. Let $T : X \to X$ be a linear operator such that*

$$\|Tx\| \leq \|x\|, \qquad \forall\, x \in X.$$

Then for every $x \in X$ the limit

$$p_x = \lim_{n\to\infty} \frac{x + Tx + \cdots + T^n x}{n+1}$$

exists. Moreover, the operator $P : X \to X$ defined by $Px = p_x$ is a continuous projection onto the linear space $\mathcal{M} = \{y \in X : Ty = y\}$.

PROOF Fix $x \in X$, and set

$$C = \overline{\operatorname{co}(\{x, Tx, T^2x, T^3x, \dots\})}.$$

C is a closed nonvoid convex set, and from the uniform convexity of X there is a unique $p_x \in C$ such that

$$\mu = \|p_x\| = \inf\big\{\|z\| : z \in C\big\}.$$

Select $\varepsilon > 0$. Then, for $p_x \in C$, there are $m \in \mathbb{N}$ and nonnegative constants $\alpha_0, \alpha_1, \dots, \alpha_m$ with $\sum_{j=0}^m \alpha_j = 1$ such that, setting

$$z = \sum_{j=0}^m \alpha_j T^j x$$

there holds

$$\|p_x - z\| < \varepsilon.$$

In particular, for every $n \in \mathbb{N}$,

$$\left\| \frac{z + Tz + \cdots + T^n z}{n+1} \right\| \leq \|z\| \leq \mu + \varepsilon.$$

Notice that

$$\begin{aligned} z + Tz + \cdots + T^n z \quad &= \quad (\alpha_0 x + \cdots + \alpha_m T^m x) + (\alpha_0 Tx + \cdots + \alpha_m T^{m+1} x) \\ &\quad + \cdots + (\alpha_0 T^n x + \cdots + \alpha_m T^{m+n} x). \end{aligned}$$

Thus, assuming $n \gg m$, we get

$$z + Tz + \cdots + T^n z = x + Tx + \cdots + T^n x + r$$

where

$$\begin{aligned} r \quad = \quad & (\alpha_0 - 1)x + \cdots + (\alpha_0 + \alpha_1 + \cdots + \alpha_{m-1} - 1)T^{m-1} x \\ & + (1 - \alpha_0)T^{1+n} x + \cdots + (1 - \alpha_0 - \alpha_1 - \cdots - \alpha_{m-1})T^{m+n} x. \end{aligned}$$

Therefore

$$\frac{x + Tx + \cdots + T^n x}{n+1} = \frac{z + Tz + \cdots + T^n z}{n+1} - \frac{r}{n+1}.$$

Since

$$\left\| \frac{r}{n+1} \right\| \leq \frac{2m\|x\|}{n+1}$$

upon choosing n enough large such that $2m\|x\| < \varepsilon(n+1)$ we have

$$\left\| \frac{x + Tx + \cdots + T^n x}{n+1} \right\| \leq \left\| \frac{z + Tz + \cdots + T^n z}{n+1} \right\| + \left\| \frac{r}{n+1} \right\| \leq \mu + 2\varepsilon.$$

On the other hand, it must be

$$\left\| \frac{x + Tx + \cdots + T^n x}{n+1} \right\| \geq \mu.$$

Then we conclude that

$$\lim_{n \to \infty} \left\| \frac{x + Tx + \cdots + T^n x}{n+1} \right\| = \mu.$$

This says that the above is a minimizing sequence in C, and due to the uniform convexity of X, we gain the convergence

$$\lim_{n \to \infty} \frac{x + Tx + \cdots + T^n x}{n+1} = p_x.$$

We are left to show that the operator $Px = p_x$ is a continuous projection onto $\mathcal{M}$. Indeed, it is apparent that if $x \in \mathcal{M}$ then $p_x = x$. In general,

$$Tp_x = \lim_{n\to\infty} \frac{Tx + T^2x + \cdots + T^{n+1}x}{n+1} = p_x + \lim_{n\to\infty} \frac{T^{n+1}x - x}{n+1} = p_x.$$

Finally, $P^2x = PPx = Pp_x = p_x = Px$. The continuity is ensured by the relation $\|p_x\| \leq \|x\|$. $\diamond$

When X is a Hilbert space, P is actually an orthogonal projection. This follows from the next proposition.

Proposition *Let H be a Hilbert space, $P = P^2 : H \to H$ a bounded linear operator with $\|P\| \leq 1$. Then P is an orthogonal projection.*

PROOF Since P is continuous, $\mathrm{ran}(P)$ is closed. Let then E be the orthogonal projection having range $\mathrm{ran}(P)$. Then

$$P = E + P(\mathbb{I} - E).$$

Let now $x \in \mathrm{ran}(P)^\perp$. For any $\varepsilon > 0$ we have

$$\|P(Px + \varepsilon x)\| \leq \|Px + \varepsilon x\|$$

which implies that

$$\|Px\|^2 \leq \frac{\varepsilon}{2+\varepsilon}\|x\|^2.$$

Hence $Px = 0$, and the equality $P = E$ holds. $\diamond$

The role played by uniform convexity in Theorem 1.13 is essential, as the following example shows.

Example Let $X = \ell^\infty$, and let $T \in L(X)$ defined by

$$Tx = (0, x_0, x_1, \ldots), \qquad \text{for } x = (x_0, x_1, x_2, \ldots) \in X.$$

Then T has a unique fixed point, namely, the zero element of X. Nonetheless, if $y = (1, 1, 1, \ldots)$, for every $n \in \mathbb{N}$ we have

$$\left\| \frac{y + Ty + \cdots + T^n y}{n+1} \right\| = \frac{\|(1, 2, \ldots, n, n+1, n+1, \ldots)\|}{n+1} = 1.$$

The Brouwer fixed point theorem

Let $\mathbb{D}^n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$. A subset $E \subset \mathbb{D}^n$ is called a *retract* of $\mathbb{D}^n$ if there exists a continuous map $r : \mathbb{D}^n \to E$ (called *retraction*) such that $r(x) = x$ for every $x \in E$.

1.14 Lemma *The set* $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ *is not a retract of* $\mathbb{D}^n$.

The lemma can be easily proved by means of algebraic topology tools. Indeed, a retraction r induces a homomorphism $r_* : H_{n-1}(\mathbb{D}^n) \to H_{n-1}(\mathbb{S}^{n-1})$, where H_{n-1} denotes the $(n-1)$-dimensional *homology group* (see, e.g., [8]). The natural injection $j : \mathbb{S}^{n-1} \to \mathbb{D}^n$ induces in turn a homomorphism $j_* : H_{n-1}(\mathbb{S}^{n-1}) \to H_{n-1}(\mathbb{D}^n)$, and the composition rj is the identity map on $\mathbb{S}^{n-1}$. Hence $(rj)_* = r_* j_*$ is the identity map on $H_{n-1}(\mathbb{S}^{n-1})$. But since $H_{n-1}(\mathbb{D}^n) = 0$, j_* is the null map. On the other hand, $H_{n-1}(\mathbb{S}^{n-1}) = \mathbb{Z}$ if $n \neq 1$, and $H_0(\mathbb{S}^0) = \mathbb{Z} \oplus \mathbb{Z}$, leading to a contradiction.

The analytic proof reported below is less evident, and makes use of exterior forms. Moreover, it provides a weaker result, namely, it shows that there exist no retraction of class C^2 from $\mathbb{D}^n$ to $\mathbb{S}^{n-1}$. This will be however enough for our scopes.

PROOF Associate to a C^2 function $h : \mathbb{D}^n \to \mathbb{D}^n$ the exterior form

$$\omega_h = h_1 \, dh_2 \wedge \cdots \wedge dh_n.$$

The Stokes theorem (cf. [12], Ch.10) entails

$$\mathcal{D}_h := \int_{\mathbb{S}^{n-1}} \omega_h = \int_{\mathbb{D}^n} d\omega_h = \int_{\mathbb{D}^n} dh_1 \wedge \cdots \wedge dh_n = \int_{\mathbb{D}^n} \det[J_h(x)] \, dx$$

where $J_h(x)$ denotes the $(n \times n)$-Jacobian matrix of h at x. Assume now that there is a retraction r of class C^2 from $\mathbb{D}^n$ to $\mathbb{S}^{n-1}$. From the above formula, we see that $\mathcal{D}_r$ is determined only by the values of r on $\mathbb{S}^{n-1}$. But $r_{|\mathbb{S}^{n-1}} = i_{|\mathbb{S}^{n-1}}$, where $i : \mathbb{D}^n \to \mathbb{D}^n$ is the identity map. Thus $\mathcal{D}_r = \mathcal{D}_i = \mathrm{vol}\,(\mathbb{D}^n)$. On the other hand $\|r\| \equiv 1$, and this implies that the vector $J_r(x) r(x)$ is null for every $x \in \mathbb{D}^n$. So 0 is an eigenvalue of $J_r(x)$ for every $x \in \mathbb{D}^n$, and therefore $\det[J_r] \equiv 0$ which implies $\mathcal{D}_r = 0$. ◇

Remark Another way to show that $\det[J_r] \equiv 0$ is to observe that the determinant is a *null lagrangian* (for more details see, e.g., [5], Ch.8)

1.15 Theorem [Brouwer] *Let* $f : \mathbb{D}^n \to \mathbb{D}^n$ *be a continuous function. Then* f *has a fixed point* $\bar{x} \in \mathbb{D}^n$.

PROOF Since we want to rely on the analytic proof, let $f : \mathbb{D}^n \to \mathbb{D}^n$ be of class C^2. If f had no fixed point, then

$$r(x) = t(x) f(x) + (1 - t(x)) x$$

where

$$t(x) = \frac{\|x\|^2 - \langle x, f(x) \rangle - \sqrt{(\|x\|^2 - \langle x, f(x) \rangle)^2 + (1 - \|x\|^2)\|x - f(x)\|^2}}{\|x - f(x)\|^2}$$

is a retraction of class C^2 from $\mathbb{D}^n$ to $\mathbb{S}^{n-1}$, against the conclusion of Lemma 1.14. Graphically, $r(x)$ is the intersection with $\mathbb{S}^{n-1}$ of the line obtained extending the segment connecting $f(x)$ to x. Hence such an f has a fixed point. Finally, let $f : \mathbb{D}^n \to \mathbb{D}^n$ be continuous. Appealing to the Stone-Weierstrass theorem, we find a sequence $f_j : \mathbb{D}^n \to \mathbb{D}^n$ of functions of class C^2 converging uniformly to f on $\mathbb{D}^n$. Denote $\bar{x}_j$ the fixed point of f_j. Then there is $\bar{x} \in \mathbb{D}^n$ such that, up to a subsequence, $\bar{x}_j \to \bar{x}$. Therefore,

$$\|f(\bar{x}) - \bar{x}\| \leq \|f(\bar{x}) - f(\bar{x}_j)\| + \|f(\bar{x}_j) - f_j(\bar{x}_j)\| + \|\bar{x}_j - \bar{x}\| \longrightarrow 0$$

as $j \to \infty$, which yields $f(\bar{x}) = \bar{x}$. $\diamond$

Remark An alternative approach to prove Theorem 1.15 makes use of the concept of *topological degree* (see, e.g., [2], Ch.1).

The Schauder-Tychonoff fixed point theorem

We first extend Theorem 1.15 to a more general situation.

1.16 Lemma *Let K be a nonvoid compact convex subset of a finite dimensional real Banach space X. Then every continuous function $f : K \to K$ has a fixed point $\bar{x} \in K$.*

PROOF Since X is homeomorphic to $\mathbb{R}^n$ for some $n \in \mathbb{N}$, we assume without loss of generality $X = \mathbb{R}^n$. Also, we can assume $K \subset \mathbb{D}^n$. For every $x \in \mathbb{D}^n$, let $p(x) \in K$ be the unique point of minimum norm of the set $x - K$. Notice that $p(x) = x$ for every $x \in K$. Moreover, p is continuous on $\mathbb{D}^n$. Indeed, given $x_n, x \in \mathbb{D}^n$, with $x_n \to x$,

$$\|x - p(x)\| \leq \|x - p(x_n)\| \leq \|x - x_n\| + \inf_{k \in K} \|x_n - k\| \longrightarrow \|x - p(x)\|$$

as $n \to \infty$. Thus $x - p(x_n)$ is a minimizing sequence as $x_n \to x$ in $x - K$, and this implies the convergence $p(x_n) \to p(x)$. Define now $g(x) = f(p(x))$. Then g maps continuously $\mathbb{D}^n$ onto K. From Theorem 1.15 there is $\bar{x} \in K$ such that $g(\bar{x}) = \bar{x} = f(\bar{x})$. $\diamond$

As an immediate application, consider Example 1.1. If there is a compact and convex set $K \subset \mathbb{R}^n$ such that $h(K) \subset K$, then h has a fixed point $\bar{x} \in K$.

Other quite direct applications are the Frobenius theorem and the fundamental theorem of algebra.

Theorem [Frobenius] *Let $\mathbb{A}$ be a $n \times n$ matrix with strictly positive entries. Then $\mathbb{A}$ has a strictly positive eigenvalue.*

PROOF The matrix $\mathbb{A}$ can be viewed as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. Introduce the compact convex set

$$K = \Big\{ x \in \mathbb{R}^n : \sum_{j=1}^n x_j = 1, \;\; x_j \geq 0 \;\text{ for }\; j = 1, \ldots, n \Big\}$$

and define $f(x) = \mathbb{A}x/\|\mathbb{A}x\|_1$ (where $\|\cdot\|_1$ is the euclidean 1-norm). Notice that if $x \in K$, then all the entries of x are nonnegative and at least one is strictly positive, hence all the entries of $\mathbb{A}x$ are strictly positive. Then f is a continuous function mapping K into K, and therefore there exists $\bar{x} \in K$ such that $\mathbb{A}\bar{x} = \|\mathbb{A}\bar{x}\|_1\,\bar{x}$. ◇

Theorem [Fundamental of algebra] *Let $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ be a complex polynomial of degree $n \geq 1$. Then there exits $z_0 \in \mathbb{C}$ such that $p(z_0) = 0$.*

PROOF For our purposes, let us identify $\mathbb{C}$ with $\mathbb{R}^2$. Suppose without loss of generality $a_n = 1$. Let $r = 2 + |a_0| + \cdots + |a_{n-1}|$. Define now the continuous function $g : \mathbb{C} \to \mathbb{C}$ as

$$g(z) = \begin{cases} z - \dfrac{p(z)}{r}\, e^{i(1-n)\vartheta} & |z| \leq 1 \\ z - \dfrac{p(z)}{r}\, z^{(1-n)} & |z| > 1 \end{cases}$$

where $z = \rho e^{i\vartheta}$ with $\vartheta \in [0, 2\pi)$. Consider now the compact and convex set $C = \{z : |z| \leq r\}$. In order to apply the Brouwer fixed point theorem we need to show that $g(C) \subset C$. Indeed, if $|z| \leq 1$,

$$|g(z)| \leq |z| + \frac{|p(z)|}{r} \leq 1 + \frac{1 + |a_0| + \cdots + |a_{n-1}|}{r} \leq 2 \leq r.$$

Conversely, if $1 < |z| \leq r$ we have

$$\begin{aligned} |g(z)| &\leq \left| z - \frac{p(z)}{r z^{n-1}} \right| = \left| z - \frac{z}{r} - \frac{a_0 + a_1 z + \cdots a_{n-1} z^{n-1}}{r z^{n-1}} \right| \\ &\leq r - 1 + \frac{|a_0| + \cdots + |a_{n-1}|}{r} \leq r - 1 + \frac{r-2}{r} \leq r. \end{aligned}$$

Hence C is invariant for g, and so g has a fixed point $z_0 \in C$, which is clearly a root of p. ◇

1.17 Partition of the unity Suppose $V_1, \ldots, V_n$ are open subsets of a locally compact Hausdorff space X, $K \subset X$ is compact, and

$$K \subset V_1 \cup \cdots \cup V_n.$$

Then for every $j = 1, \ldots, n$ there exists $\varphi_j \in C(X)$, $0 \leq \varphi_j \leq 1$, supported on V_j such that

$$\varphi_1(x) + \cdots + \varphi_n(x) = 1, \qquad \forall\, x \in K.$$

The collection $\varphi_1, \ldots, \varphi_n$ is said to be a *partition of the unity* for K subordinate to the open cover $\{V_1, \ldots, V_n\}$.

The existence of a partition of the unity is a straightforward consequence of the Urysohn lemma (see for instance [7]). We are often interested to find partitions of the unity for a compact set $K \subset X$ whose members are continuous functions defined on K. Clearly, in this case X need not be locally compact.

1.18 Theorem [Schauder-Tychonoff] *Let X be a locally convex space, $K \subset X$ nonvoid and convex, $K_0 \subset K$, K_0 compact. Given a continuous map $f : K \to K_0$, there exists $\bar{x} \in K_0$ such that $f(\bar{x}) = \bar{x}$.*

PROOF Denote by $\mathcal{B}$ the local base for the topology of X generated by the separating family of seminorms $\mathcal{P}$ on X. Given $U \in \mathcal{B}$, from the compactness of K_0, there exist $x_1, \dots, x_n \in K_0$ such that

$$K_0 \subset \bigcup_{j=1}^{n} (x_j + U).$$

Let $\varphi_1, \dots, \varphi_n \in C(K_0)$ be a partition of the unity for K_0 subordinate to the open cover $\{x_j + U\}$, and define

$$f_U(x) = \sum_{j=1}^{n} \varphi_j(f(x)) x_j, \qquad \forall\, x \in K.$$

then

$$f_U(K) \subset K_U := \mathrm{co}(\{x_1, \dots, x_n\}) \subset K$$

and Lemma 1.16 yields the existence of $x_U \in K_U$ such that $f_U(x_U) = x_U$. Then

$$x_U - f(x_U) = f_U(x_U) - f(x_U) = \sum_{j=1}^{n} \varphi_j(f(x_U))(x_j - f(x_U)) \in U \tag{2}$$

for $\varphi_j(f(x_U)) = 0$ whenever $x_j - f(x_U) \notin U$. Appealing again to the compactness of K_0, there exists

$$\bar{x} \in \bigcap_{W \in \mathcal{B}} \overline{\big\{f(x_U) : U \in \mathcal{B},\, U \subset W\big\}} \subset K_0. \tag{3}$$

Select now $p \in \mathcal{P}$ and $\varepsilon > 0$, and let

$$V = \big\{x \in X : p(x) < \varepsilon\big\} \in \mathcal{B}.$$

Since f is continuous on K, there is $W \in \mathcal{B}$, $W \subset V$, such that

$$f(x) - f(\bar{x}) \in V$$

whenever $x - \bar{x} \in 2W$, $x \in K$. Moreover, by (3), there exists $U \in \mathcal{B}$, $U \subset W$, such that

$$\bar{x} - f(x_U) \in W \subset V. \tag{4}$$

Collecting (2) and (4) we get

$$x_U - \bar{x} = x_U - f(x_U) + f(x_U) - \bar{x} \in U + W \subset W + W = 2W$$

which yields

$$f(x_U) - f(\bar{x}) \in V. \tag{5}$$

Hence (4)-(5) entail

$$p(\bar{x} - f(\bar{x})) \leq p(\bar{x} - f(x_U)) + p(f(x_U) - f(\bar{x})) < 2\varepsilon.$$

Being p and ε arbitrary, we conclude that $p(\bar{x} - f(\bar{x})) = 0$ for every $p \in \mathcal{P}$, which implies the equality $f(\bar{x}) = \bar{x}$. ◇

The following two propositions deal with the existence of zeros of maps on Banach spaces. For a Banach space X and $r > 0$, let $B_r = \overline{B}_X(0, r)$, and consider a continuous map $g : B_r \to X$, such that $g(B_r)$ is relatively compact.

Proposition *Let $g(x) \notin \{\lambda x : \lambda > 0\}$ for every $x \in \partial B_r$. Then there is $x_0 \in B_r$ such that $g(x_0) = 0$.*

PROOF If not so, the function $f(x) = rg(x)/\|g(x)\|$ is continuous from B_r to B_r and $f(B_r)$ is relatively compact. From Theorem 1.18, f admits a fixed point $\bar{x} \in B_r$. Then $g(\bar{x}) = \|g(\bar{x})\|\bar{x}/r$ with $\|\bar{x}\| = r$, against the hypotheses. ◇

Proposition *Assume that for every $x \in \partial B_r$ there exists $\Lambda_x \in X^*$, $\Lambda_x x = 1$, such that $\Lambda_x g(x) \geq 0$. Then there is $x_0 \in B_r$ such that $g(x_0) = 0$.*

PROOF Define $f(x) = -rg(x)/\|g(x)\|$. If g has no zeros, reasoning as above, f has a fixed point $\bar{x} \in B_r$, and the relation $-g(\bar{x}) = \|g(\bar{x})\|\bar{x}/r$, with $\|\bar{x}\| = r$, holds. Taking $\Lambda \in X^*$ such that $\Lambda\bar{x} = 1$, we obtain $\Lambda g(\bar{x}) = -\|g(\bar{x})\|/r < 0$. Contradiction. ◇

Let us see an interesting corollary about surjectivity of maps from $\mathbb{R}^n$ to $\mathbb{R}^n$, that extends a well-known result for matrices to the more general situation of continuous maps.

Corollary *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous function satisfying*

$$\lim_{\|x\| \to \infty} \frac{\langle f(x), x \rangle}{\|x\|} = \infty.$$

Then $f(\mathbb{R}^n) = \mathbb{R}^n$.

PROOF Fix $y_0 \in \mathbb{R}^n$, and set $g(x) = f(x) - y_0$. Then, for $r > 0$ big enough,

$$\langle g(x), x/\|x\| \rangle \geq 0, \qquad \forall\, x \in \partial B_r.$$

Hence for every $x \in \partial B_r$, the functional $\Lambda_x := \langle \cdot, x/\|x\| \rangle$ fulfills the hypotheses of the last proposition. Therefore there is $x_0 \in B_r$ such that $g(x_0) = 0$, that is, $f(x_0) = y_0$. ◇

Remark These kind of conditions concerning the behavior of g on ∂B_r are known as *Leray-Schauder boundary conditions.* The above results can be generalized to continuous functions defined on the closure of open subsets of X with values in X satisfying certain compactness properties (see for instance [2], pp.204–205, and [6], Ch.5.3).

In the applications it is somehow easier to work with functions defined on the whole space X, and rather ask more restrictive conditions, such as compactness, on the maps.

1.19 Definition Let X, Y be Banach spaces, and $C \subset X$. A map $f : C \to Y$ is said to be *compact* provided it transforms bounded sets into relatively compact sets. If $f \in L(X, Y)$ it is the same as saying that the image of the unit closed ball under f is relatively compact. In terms of sequences, f is compact if for every bounded sequence x_n, the sequence $f(x_n)$ has a convergent subsequence.

1.20 Theorem [Schaefer] *Let X be a Banach space, $f : X \to X$ continuous and compact. Assume further that the set*

$$F = \big\{ x \in X \,:\, x = \lambda f(x) \quad \text{for some} \quad \lambda \in [0,1] \big\}$$

is bounded. Then f has a fixed point.

Remark Theorem 1.20 holds if we can prove *a priori* estimates on the set of all the *possible* fixed point of λf. This technique is typical in partial differential equations, where one proves estimates on a possible solution to a certain equation, and then, in force of those estimates, concludes that such solution actually exists.

PROOF Let $r > \sup_{x \in F} \|x\|$, and define the map

$$g(x) = \begin{cases} f(x) & \text{if} \quad \|f(x)\| \le 2r \\ \dfrac{2r f(x)}{\|f(x)\|} & \text{if} \quad \|f(x)\| > 2r. \end{cases}$$

Then $g : \overline{B}_X(0, 2r) \to \overline{B}_X(0, 2r)$ is continuous and compact. Theorem 1.18 yields the existence of $x_0 \in \overline{B}_X(0, 2r)$ for which $g(x_0) = x_0$. Notice that $\|f(x_0)\| \le 2r$, for otherwise

$$x_0 = \lambda_0 f(x_0) \qquad \text{with} \qquad \lambda_0 = \frac{2r}{\|f(x_0)\|} < 1$$

which forces $\|x_0\| = 2r$, against the fact that $x_0 \in F$. Hence we get that $g(x_0) = f(x_0) = x_0$. $\diamond$

1.21 Theorem [Krasnoselskii] *Let X be a Banach space, $C \subset X$ nonvoid, closed and convex. Let $f, g : C \to X$ be such that*

(a) $f(x_1) + g(x_2) \in C, \quad \forall\, x_1, x_2 \in C;$

(*b*) *f is continuous and compact;*

(*c*) *g is a contraction (from C into X).*

Then there is $\bar{x} \in C$ such that $f(\bar{x}) + g(\bar{x}) = \bar{x}$.

PROOF Notice first that $\mathbb{I}-g$ maps homeomorphically C onto $(\mathbb{I}-g)(C)$. Indeed, $\mathbb{I}-g$ is continuous and

$$\|(\mathbb{I}-g)(x_1) - (\mathbb{I}-g)(x_2)\| \geq \|x_1 - x_2\| - \|g(x_1) - g(x_2)\| \geq (1-\lambda)\|x_1 - x_2\|$$

($\lambda < 1$ is the Lipschitz constant of g) so $(\mathbb{I}-g)^{-1}$ is continuous. For any $y \in C$, the map $x \mapsto f(y) + g(x)$ is a contraction on C, hence by Theorem 1.3 there is a unique $z = z(y) \in C$ such that $z = f(y) + g(z)$. Thus $z = (\mathbb{I}-g)^{-1}(f(y)) \in C$. On the other hand, the map $(\mathbb{I}-g)^{-1} \circ f$ is continuous and compact from C to C, being the composition of a continuous map with a continuous and compact map. Then Theorem 1.18 entails the existence of $\bar{x} \in C$ such that $(\mathbb{I}-g)^{-1}(f(\bar{x})) = \bar{x}$, that is, $f(\bar{x}) + g(\bar{x}) = \bar{x}$. $\diamond$

In general, it is not possible to extend Theorem 1.18 to noncompact settings. This fact was already envisaged in our previous discussion about non-expansive maps. Let us recall another famous example in Hilbert spaces.

Example [Kakutani] Consider the Hilbert space ℓ^2. For a fixed $\varepsilon \in (0,1]$, let $f_\varepsilon : \overline{B}_{\ell^2}(0,1) \to \overline{B}_{\ell^2}(0,1)$ be given by

$$f_\varepsilon(x) = (\varepsilon(1 - \|x\|), x_0, x_1, \ldots), \qquad \text{for } x = (x_0, x_1, x_2, \ldots) \in \ell^2.$$

Then f_ε has no fixed points in $\overline{B}_{\ell^2}(0,1)$, but it is Lipschitz continuous with Lipschitz constant slightly greater than 1. Indeed,

$$\|f_\varepsilon(x) - f_\varepsilon(y)\| \leq \sqrt{1+\varepsilon^2}\,\|x - y\|$$

for all $x, y \in \overline{B}_{\ell^2}(0,1)$.

It is therefore a natural question to ask when a continuous map $f : C \to C$, where C is a closed, bounded and convex subset of a Banach space X, admits fixed points. We know from Theorem 1.18 that if X is finite-dimensional the answer is positive, since finite-dimensional Banach spaces have the Heine-Borel property. The analogous result in infinite-dimensional Banach spaces turns out to be false, without a compactness assumption.

We first report a characterization of noncompact closed bounded sets.

Lemma *Let X be a Banach space, $C \subset X$ a closed, bounded noncompact set. Then there are $\varepsilon > 0$ and a sequence x_n of elements of C such that*

$$\operatorname{dist}\big(x_{n+1}, \operatorname{span}(\{x_0, \ldots, x_n\})\big) \geq \varepsilon. \tag{6}$$

PROOF Clearly, X has to be infinite-dimensional, otherwise no such C exists. We first show that there is $\varepsilon > 0$ such that, for any finite set $F \subset X$,

$$C \setminus [\text{span}(F) + B_X(0, \varepsilon)] \neq \emptyset.$$

If not, for every $\varepsilon > 0$ we find a finite set $F \subset X$ such that $C \subset \text{span}(F) + B_X(0, \varepsilon)$. Since C is bounded, $C \subset B_X(0, r)$ for some $r > 0$. Therefore,

$$C \subset [\text{span}(F) + B_X(0, \varepsilon)] \cap B_X(0, r) \subset [\text{span}(F) \cap B_X(0, r + \varepsilon)] + B_X(0, \varepsilon).$$

But $\text{span}(F) \cap B_X(0, r + \varepsilon)$ is totally bounded, and so it admits a finite cover of balls of radius ε, which in turn implies that C admits a finite cover of balls of radius 2ε, that is, C is totally bounded (hence compact), contradicting the hypotheses.

To construct the required sequence x_n, we proceed inductively. First we select an arbitrary $x_0 \in C$. If we are given $x_0, \ldots, x_{n+1}$ satisfying (6), we choose $x_{n+2} \in C \setminus [\text{span}(\{x_0, \ldots, x_{n+1}\}) + B_X(0, \varepsilon)]$. ◇

1.22 Theorem [Klee] *Let X be an infinite-dimensional Banach space, $C \subset X$ a closed, bounded, convex noncompact set. Then there exists a continuous map $f : C \to C$ which is fixed point free.*

PROOF Let x_n be a sequence in C satisfying (6). Without loss of generality, we may assume that $0 \in C$ and $\|x_0\| \geq \varepsilon$. We construct a piecewise linear curve joining the points $x_0, x_1, x_2, \ldots$, by setting

$$\Gamma = \bigcup_{n=0}^{\infty} [x_n, x_{n+1}], \qquad \text{where } [x_n, x_{n+1}] = \text{co}(\{x_n, x_{n+1}\}).$$

Then $\Gamma \subset C$ is closed and can be parameterized by a function $\gamma : [0, \infty) \to \Gamma$, given by

$$\gamma(t) = (1 - s)x_n + s x_{n+1}$$

where $n = [t]$ and $s = t - n$. By means of (6), γ is one-to-one, onto and open, since for every open set $O \subset [0, \infty)$, $\gamma(O)$ contains the intersection of Γ with an open ball of X of radius ν, for some $\nu < \varepsilon$. Hence its inverse $\gamma^{-1} : \Gamma \to [0, \infty)$ is a continuous function. Applying a slightly modified version of the Tietze extension theorem (see, e.g., [7]), we can extend γ^{-1} to a continuous function $g : C \to [0, \infty)$. Now a fixed point free continuous map $f : C \to C$ can be defined as

$$f(x) = \gamma(g(x) + 1).$$

Indeed, if $f(\bar{x}) = \bar{x}$, then $\bar{x} \in \Gamma$. Hence $\gamma(\gamma^{-1}(\bar{x}) + 1) = \gamma(\gamma^{-1}(\bar{x}))$, against the injectivity of γ. ◇

Remark Notice however that Klee's map $f : C \to C$ cannot be uniformly continuous, since the restriction of g on Γ is easily seen to be uniformly continuous, whereas $g \circ f$ maps the bounded convex set C onto $[1, \infty)$.

In particular, Theorem 1.22 says that in an infinite-dimensional Banach space X there is a continuous map $f : \overline{B}_X(0,1) \to \overline{B}_X(0,1)$ without fixed points. This allow us to provide another interesting characterization of infinite-dimensional Banach spaces.

Corollary *Let X be an infinite-dimensional Banach space. Then the closed unit sphere is a retract of the closed unit ball.*

PROOF Let $f : \overline{B}_X(0,1) \to \overline{B}_X(0,1)$ be a continuous fixed point free map. Extend the map to the doubled ball $\overline{B}_X(0,2)$ by defining

$$f_1(x) = \begin{cases} f(x) & \text{if} \quad \|x\| \le 1 \\ (2 - \|x\|) f(x/\|x\|) & \text{if} \quad 1 < \|x\| \le 2 \end{cases}$$

and consider the new map $f_2 : \overline{B}_X(0,1) \to \overline{B}_X(0,1)$ as

$$f_2(x) = \frac{1}{2} f_1(2x).$$

Observe that f_2 is fixed point free, and for all $x \in \partial\overline{B}_X(0,1)$ we have $f_2(x) = 0$. Then the function $r : \overline{B}_X(0,1) \to \partial\overline{B}_X(0,1)$ given by

$$r(x) = \frac{x - f_2(x)}{\|x - f_2(x)\|}$$

is the desired retraction. $\diamond$

The Markov-Kakutani theorem

The following theorem is concerned with common fixed points of a family of linear maps.

1.23 Theorem [Markov-Kakutani] *Let X be a locally convex space, and let $K \subset X$ be a nonvoid, compact and convex set. Assume $\mathcal{G}$ is a family of bounded linear operators from X into X such that*

(a) *$\mathcal{G}$ is abelian, that is, $TS = ST$ for every $T, S \in \mathcal{G}$;*

(b) *$TK \subset K$ for every $T \in \mathcal{G}$.*

Then there exists $\bar{x} \in K$ such that $T\bar{x} = \bar{x}$ for every $T \in \mathcal{G}$.

PROOF For any $T \in \mathcal{G}$ and any $n \in \mathbb{N}$, define the operator

$$T_n = \frac{\mathbb{I} + T + \cdots + T^n}{n+1}.$$

Notice that (b) and the convexity of K imply that

$$T_n K \subset K.$$

Given $T^{(1)}, \dots, T^{(k)} \in \mathcal{G}$ and $n_1, \dots, n_k \in \mathbb{N}$, it follows that

$$\bigcap_{j=1}^{k} T_{n_j}^{(j)} K \neq \emptyset.$$

Indeed, from (a), for any $T, S \in \mathcal{G}$ and any $n, m \in \mathbb{N}$,

$$T_n K \cap S_m K \supset T_n S_m K = S_m T_n K \neq \emptyset.$$

Invoking the compactness of K, the finite intersection property yields

$$F = \bigcap_{T \in \mathcal{G},\, n \in \mathbb{N}} T_n K \neq \emptyset.$$

We claim that every $x \in F$ is a fixed point for all $T \in \mathcal{G}$. Let then $\bar{x} \in F$, and select any $T \in \mathcal{G}$. Then for every $n \in \mathbb{N}$ there is $y = y(n) \in K$ such that $\bar{x} = T_n y$. Hence

$$\begin{aligned} T\bar{x} - \bar{x} &= \frac{Ty + T^2 y + \cdots + T^{n+1} y}{n+1} - \frac{y + Ty + \cdots + T^n y}{n+1} \\ &= \frac{T^{n+1} y - y}{n+1} \in \frac{1}{n+1}\big(K - K\big). \end{aligned}$$

Thus

$$T\bar{x} - \bar{x} \in \bigcap_{n \in \mathbb{N}} \frac{1}{n+1}\big(K - K\big).$$

The set $K - K$ is clearly compact, being the image of $K \times K$ under the continuous map $\Phi(x_1, x_2) = x_1 - x_2$. Let then p be a seminorm on X. Then, for any $\varepsilon > 0$, there is $n \in \mathbb{N}$ such that

$$\frac{1}{n+1}\big(K - K\big) \subset \big\{x \in X : p(x) < \varepsilon\big\}.$$

We conclude that $p(T\bar{x} - \bar{x}) = 0$ for every seminorm on X, which entails the equality $T\bar{x} = \bar{x}$. $\diamond$

Remark With no changes in the proof, the results holds more generally if $\mathcal{G}$ is an abelian family of continuous affine maps from K to K. A map $f : C \to C$, C convex, is said to be *affine* if for all $x_1, x_2 \in C$ and $t \in [0, 1]$

$$f(tx_1 + (1-t)x_2) = tf(x_1) + (1-t)f(x_2).$$

The Kakutani-Ky Fan theorem

Along the section, let X be locally convex space.

Definition Let $C \subset X$ be a convex set. A function $f : C \to (-\infty, \infty]$ is said to be *convex* if

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

for all $\forall\, x_1, x_2 \in X$ and every $\lambda \in [0,1]$. A function $g : C \to [-\infty, \infty)$ is said to be *concave* if $-g$ is convex (it is understood that $-(\infty) = -\infty$).

We now recall the well-known definition of lower and upper semicontinuous real functions.

Definition Let Y be a topological space. A function $f : Y \to (-\infty, \infty]$ is said to be *lower semicontinuous* if $f^{-1}((\alpha, \infty])$ is open for every $\alpha \in \mathbb{R}$. Similarly, a function $g : Y \to [-\infty, \infty)$ is said to be *upper semicontinuous* if $-g$ is lower semicontinuous.

It is an immediate consequence of the definition that the supremum of any collection of lower semicontinuous functions is lower semicontinuous. Moreover, if f is lower semicontinuous and Y is compact, then f attains its minimum on Y. Indeed, if it is not so, denoting $m = \inf_{y \in Y} f(y) \in [-\infty, \infty)$, the sets $f^{-1}((\alpha, \infty])$ with $\alpha > m$ form an open cover of Y that admits no finite subcovers.

The next result is the famous Ky Fan inequality.

1.24 Theorem [Ky Fan] *Let $K \subset X$ be a nonvoid, compact and convex. Let $\Phi : K \times K \to \mathbb{R}$ be map such that*

(*a*) $\Phi(\cdot, y)$ *is lower semicontinuous* $\forall\, y \in K$*;*

(*b*) $\Phi(x, \cdot)$ *is concave* $\forall\, x \in K$*.*

Then there exists $x_0 \in K$ such that

$$\sup_{y \in K} \Phi(x_0, y) \leq \sup_{y \in K} \Phi(y, y).$$

PROOF Fix $\varepsilon > 0$. In correspondence with every $x \in K$ there are $y_x \in K$ and an open neighborhood U_x of x such that

$$\Phi(z, y_x) > \sup_{y \in K} \Phi(x, y) - \varepsilon, \qquad \forall\, z \in U_x \cap K.$$

Being K compact, for some $x_1, \ldots, x_n \in K$ there holds

$$K \subset U_{x_1} \cup \cdots \cup U_{x_n}.$$

Let $\varphi_1, \ldots, \varphi_n \in C(K)$ be a partition of the unity for K subordinate to the open cover $\{U_{x_j}\}$, and define

$$f(x) = \sum_{j=1}^{n} \varphi_j(x) y_{x_j}, \qquad \forall\, x \in K.$$

The map f is clearly continuous, and

$$f(\mathrm{co}(\{y_{x_1}, \ldots, y_{x_n}\})) \subset \mathrm{co}(\{y_{x_1}, \ldots, y_{x_n}\}).$$

Hence by Lemma 1.16 f admits a fixed point $\bar{x} \in K$. Therefore,

$$\begin{aligned}
\sup_{y \in K} \Phi(y,y) &\geq \Phi(\bar{x},\bar{x}) \geq \sum_{j=1}^{n} \varphi_j(\bar{x}) \Phi(\bar{x}, y_{x_j}) \\
&\geq \sum_{j=1}^{n} \varphi_j(\bar{x}) \big(\sup_{y \in K} \Phi(x_j, y) - \varepsilon \big) \\
&\geq \inf_{x \in K} \sup_{y \in K} \Phi(x,y) - \varepsilon \\
&= \sup_{y \in K} \Phi(x_0, y) - \varepsilon
\end{aligned}$$

for some $x_0 \in K$. Letting $\varepsilon \to 0$ we get the desired conclusion. $\diamond$

The aim of this section is to consider a fixed point theorem for maps carrying points into sets. Let $K \subset X$, and consider a map $f : K \to 2^K := \{Y : Y \subset K\}$. A fixed point for f is a point $x \in K$ such that $x \in f(x)$.

Definition The map f is *upper semicontinuous* if for every $x \in K$ and every open set $U \supset f(x)$, there exists a neighborhood V of x such that if $y \in V$ then $f(y) \subset U$.

Upper semicontinuous point-closed maps from a compact set K with values in 2^K can be characterized in terms of graphs.

Proposition *Let $K \subset X$ be compact, and let $f : K \to 2^K$ be such that $f(x)$ is closed for every $x \in K$. Then f is upper semicontinuous if and only if the set*

$$G(f) = \big\{(x,y) \in K \times K : y \in f(x)\big\}$$

is closed in $K \times K$.

PROOF Let f be upper semicontinuous, and $(x_0, y_0) \in (K \times K) \setminus G(f)$. Then $y_0 \notin f(x_0)$. Since K is compact, we find two disjoint open sets U_1, U_2 such that $y_0 \in U_1$ and $f(x_0) \subset U_2$. By the upper semicontinuity of f, there is an open set $V \ni x_0$ such that $f(x) \subset U_2$ for all $x \in V$. Thus the neighborhood $V \times U_1$ of (x_0, y_0) does not intersect $G(f)$, which is henceforth closed.
Conversely, let $G(f)$ be closed. Let x be an arbitrary point of K, and U be an

arbitrary open set containing $f(x)$. If f is not upper semicontinuous at x, for every neighborhood $V \ni x$ there is $y \in V$ such that $f(y) \not\subset U$. So we have

$$G(f) \cap (\overline{V} \times U^C) \neq \emptyset.$$

But $G(f)$ and $\overline{V} \times U^C$ are both compact, being closed in $K \times K$, and since the finite intersection property holds, we conclude that there is $(x_0, y_0) \in K \times K$ such that

$$(x_0, y_0) \in \bigcap_{V \ni x} G(f) \cap (\overline{V} \times U^C).$$

This implies that $x_0 = x$, and since $y_0 \in f(x_0) = f(x)$, it follows that $y_0 \in U$. Contradiction. ◇

1.25 Theorem [Kakutani-Ky Fan] *Let K be a nonvoid, compact and convex subset of a locally convex space X. Let $f : K \to 2^K$ be upper semicontinuous, such that $f(x)$ is nonvoid, convex and closed for every $x \in K$. Then f has a fixed point $\bar{x} \in K$.*

PROOF If we assume the theorem false, from the Hahn-Banach theorem for every $x \in K$ there are $\alpha_x \in \mathbb{R}$ and $\Lambda_x \in X^*$ such that

$$\sup_{z \in f(x)} \operatorname{Re} \Lambda_x z < \alpha_x < \operatorname{Re} \Lambda_x x.$$

The set

$$U_x = \big\{ z \in X : \operatorname{Re} \Lambda_x z < \alpha_x \big\}$$

is open, and contains $f(x)$. Hence there is an open neighborhood V_x of x such that $f(y) \subset U_x$ whenever $y \in V_x \cap K$. We conclude that there exists an open neighborhood $W_x \subset V_x$ of x such that

$$\operatorname{Re} \Lambda_x y > \alpha_x, \qquad \forall\, y \in W_x \cap K$$

and

$$\operatorname{Re} \Lambda_x z < \alpha_x, , \qquad \forall\, y \in W_x \cap K, \ \ \forall\, z \in f(y).$$

From the compactness of K, there exist $x_1, \dots, x_n \in K$ such that

$$K \subset W_{x_1} \cup \dots \cup W_{x_n}.$$

Let $\varphi_1, \dots, \varphi_n \in C(K)$ be a partition of the unity for K subordinate to the open cover $\{W_{x_j}\}$, and define $\Phi : K \times K \to \mathbb{R}$ as

$$\Phi(x, z) = \sum_{j=1}^{n} \varphi_j(x) \operatorname{Re} \Lambda_{x_j} x - \sum_{j=1}^{n} \varphi_j(x) \operatorname{Re} \Lambda_{x_j} z.$$

Clearly, Φ fulfills the hypotheses Theorem 1.24. Hence there is $x_0 \in K$ such that

$$\sup_{z \in K} \Phi(x_0, z) \leq \sup_{z \in K} \Phi(z, z) = 0.$$

Therefore, for $z \in f(x_0)$,

$$\sum_{j=1}^{n} \varphi_j(x_0)\alpha_{x_j} < \sum_{j=1}^{n} \varphi_j(x_0)\mathrm{Re}\,\Lambda_{x_j}x_0 \leq \sum_{j=1}^{n} \varphi_j(x_0)\mathrm{Re}\,\Lambda_{x_j}z < \sum_{j=1}^{n} \varphi_j(x_0)\alpha_{x_j}$$

which is a contradiction. Notice that in the above sums we have nonzero contributions only for those j for which $x_0 \in W_{x_j}$. ◇

Notes on Chapter 1

The construction of the approximating sequence x_n of Theorem 1.3 is known as Picard's method [*J. de Math.* **6**, 145–210 (1890)]. The abstract proof is due to Banach [*Fund. Math.* **3**, 133–181 (1922)] and Caccioppoli [*Rend. Accad. Naz. Lincei* **11**, 794–799 (1930)]. Concerning 1.6, Boyd-Wong's theorem is in *Proc. Amer. Math. Soc.* **20**, 458–464 (1969). Caristi's theorem has been proved more or less explicitly by many authors. We refer to Caristi [*Trans. Amer. Math. Soc.* **215**, 241–251 (1976)], where some applications of the theorem are also provided. Caristi's original proof involves an intricate transfinite induction argument, whereas the direct proof reported here is taken from [2], p.198. Čirič's theorem is in *Proc. Amer. Math. Soc.* **45**, 267–273 (1974). Bessaga's theorem in 1.8 appears in *Colloq. Math* **7**, 41–43 (1969). An alternative proof can be found in [2], pp.191–192. The argument we presented, indeed much simpler, is due to Peirone [personal communication] . There is an interesting related result for compact metric spaces due to Janos [*Proc. Amer. Math. Soc.* **18**, 287–289 (1967)], where an equivalent metric that makes f a contraction is constructed.

Theorem 1.9 and Theorem 1.10 are due to Nadler [*Pacif. J. Math.* **27**, 579–585 (1968)] and Fraser and Nadler [*Pacif. J. Math.* **31**, 659–667 (1969)]. Uniform convergence in Theorem 1.9 cannot in general be replaced by pointwise convergence. For instance, in every infinite-dimensional separable or reflexive Banach space there is a pointwise convergent sequence of contractions whose sequence of fixed points has no convergent subsequences (see, e.g. [6], Ch.7).

Theorem 1.11 has been proved by Kirk [*Amer. Math. Monthly* **72**, 1004–1006 (1965)], for the more general case when X is a reflexive Banach spaces and C has normal structure, and by Browder [*Proc. Nat. Acad. Sci. USA* **54**, 1041–1044 (1965)]. The proof for Hilbert spaces reported here is contained in a paper of Browder and Petryshyn [*J. Math. Anal. Appl.* **20**, 197–228 (1967)].

Theorem 1.13 is in *J. London Math. Soc.* **13**, 274–278 (1938). See also [10].

Brouwer's Theorem 1.15 is in *Math. Ann.* **71**, 97–115 (1910) (see also [3]). A different proof that requires only elementary notions of continuity and compactness has been given by Kuga [*SIAM J. Math. Anal.* **5**, 893–897 (1974)].

Schauder's original proof of Theorem 1.18 for Banach spaces can be found in *Studia Math.* **2**, 171–180 (1930), whereas Tychonoff's generalization to locally convex spaces is in *Math. Ann.* **111**, 767–776 (1935). Concerning Leray-Schauder boundary conditions, we refer the reader to the work of Leray and Schauder [*Ann. Sci. École Norm. Sup.* **51**, 45–78 (1934)]. Schaefer's Theorem 1.20 is taken from *Math. Ann.* **129**, 415–416 (1955). Theorem 1.21 is in [*Amer. Math. Soc. Transl.* **10**, 345–409 (1958)]. Theorem 1.22 has been proved by Klee [*Trans. Amer. Math. Soc.* **78**, 30–45 (1955)]. The result actually holds for metrizable locally convex spaces. In a subsequent paper [*Proc. Amer. Math. Soc.* **93**, 633–639 (1985)], Lin and Sternfeld have shown that the map f of Theorem 1.22 can be taken to be Lipschitz continuous, with Lipschitz constant greater than (but arbitrarily close to) 1.

Theorem 1.23 has been proved with different techniques by Markov [*Dokl. Akad. Nauk. SSSR* **10**, 311–314 (1936)], and Kakutani [*Proc. Imp. Acad. Tokyo* **14**, 242–245 (1938)]. The same kind of result does not hold in general for commuting nonlinear continuous maps. Indeed Boyce [*Trans. Amer. Math. Soc.* **137**, 77–92 (1969)] and Huneke [*Trans. Amer. Math. Soc.* **139**, 371–381 (1969)] have provided examples of two commuting nonlinear continuous maps of $[0, 1]$ onto itself without a common fixed point.

Theorem 1.25 for finite dimensional spaces is due to Kakutani [*Duke Math. J.* **8**, 457–459 (1941)]. The generalization to infinite dimensional spaces has been given by Ky Fan, by means of Theorem 1.24 [*Proc. Nat. Acad. Sci. USA* **88**, 121–126 (1952)].

2. SOME APPLICATIONS OF FIXED POINT THEOREMS

The implicit function theorem

2.1 Fréchet differentiability Let X, Y be (real or complex) Banach spaces, $U \subset X$, U open, $x_0 \in U$, and $f : U \to Y$.

Definition f is *Fréchet differentiable* at x_0 is there exists $T \in L(X, Y)$ and $\sigma : X \to Y$, with

$$\frac{\|\sigma(x)\|_Y}{\|x\|_X} \longrightarrow 0 \qquad \text{uniformly as} \qquad \|x\|_X \to 0$$

such that

$$f(x) - f(x_0) = T(x - x_0) + \sigma(x - x_0), \qquad \forall\, x \in U.$$

The operator T is called the *Fréchet derivative* of f at x_0, and is denoted by $f'(x_0)$. The function f is said to be Fréchet differentiable in U if it is Fréchet differentiable at every $x_0 \in U$.

It is straightforward to verify the Fréchet derivative at one point, if it exists, is unique.

2.2 Lemma *Let X, Y be Banach spaces, let $f : B_X(0, r) \to Y$ be Fréchet differentiable and $\|f'(x)\|_{L(X,Y)} \leq \lambda$ for every $x \in B_X(0, r)$ and some $\lambda \geq 0$. Then f is Lipschitz continuous with Lipschitz constant less than or equal to λ.*

PROOF Let $x_1, x_2 \in B_X(0, r)$. By the Hahn-Banach theorem, there is $\Lambda \in Y^*$ of unit norm such that

$$\|f(x_1) - f(x_2)\|_Y = |\Lambda(f(x_1) - f(x_2))|.$$

For $t \in [0, 1]$ set

$$\Phi(t) = \Lambda f(tx_1 + (1 - t)x_2).$$

Applying the Lagrange mean value theorem to Φ, there is $\tau \in (0, 1)$ such that

$$|\Lambda f(x_1) - \Lambda f(x_2)| = |\Phi(1) - \Phi(0)| \leq |\Phi'(\tau)| = |\Lambda f'(\tau x_1 + (1 - \tau)x_2)(x_1 - x_2)|$$

(the chain holds as in the classical case; if X is real equality holds). Hence

$$\|f(x_1) - f(x_2)\|_Y \leq \|f'(\tau x_1 + (1-\tau)x_2)(x_1 - x_2)\|_Y \leq \lambda\|x_1 - x_2\|_X$$

as claimed. $\diamond$

2.3 Given two Banach spaces X and Y, the vector space $X \times Y$ is a Banach space with any of the (equivalent) euclidean norms

$$\|(x,y)\|_p = \Big(\|x\|_X^p + \|y\|_Y^p\Big)^{1/p}, \qquad \|(x,y)\|_\infty = \max\big\{\|x\|_X, \|y\|_Y\big\} \qquad (p \geq 1).$$

In the sequel, we will always use the ∞-norm, so that

$$B_{X\times Y}((x_0,y_0),r) = B_X(x_0,r) \times B_Y(y_0,r).$$

For X, Y, Z Banach spaces, given $T \in L(X,Z)$ and $S \in L(Y,Z)$, the operator $R : X \times Y \to Z$ defined by

$$R(x,y) = Tx + Sy$$

belongs to $L(X \times Y, Z)$. Conversely, any $R \in L(X \times Y, Z)$, has the above representation with $Tx = R(x,0)$ and $Sy = R(0,y)$. It is then immediate to see that $L(X,Z) \times L(Y,Z)$ and $L(X \times Y, Z)$ are isomorphic Banach spaces. Given then $f : U \subset X \times Y \to Z$, U open, f Fréchet differentiable at $u_0 = (x_0, y_0) \in U$, one easily checks that the partial derivatives $D_x f(u_0)$ and $D_y f(u_0)$ exist (that is, the Fréchet derivatives of $f(\cdot, y_0) : X \to Z$ in x_0 and of $f(x_0, \cdot) : Y \to Z$ in y_0, respectively), and

$$f'(u_0)(x,y) = D_x f(u_0)(x) + D_y f(u_0)(y).$$

2.4 Theorem [Dini] *Let X, Y, Z be Banach spaces, $U \subset X \times Y$ be an open set, $u_0 = (x_0, y_0) \in U$, and $F : U \to Z$. Assume that*

(*a*) *F is continuous and $F(u_0) = 0$;*

(*b*) *$D_y F(u)$ exists for every $u = (x,y) \in U$;*

(*c*) *$D_y F$ is continuous at u_0 and $D_y F(u_0)$ is invertible.*

Then there exists $\alpha, \beta > 0$ for which $\overline{B}_X(x_0,\alpha) \times \overline{B}_Y(y_0,\beta) \subset U$ and a unique continuous function $f : \overline{B}_X(x_0,\alpha) \to \overline{B}_Y(y_0,\beta)$ such that the relation

$$F(x,y) = 0 \qquad \Longleftrightarrow \qquad y = f(x)$$

holds for all $(x,y) \in \overline{B}_X(x_0,\alpha) \times \overline{B}_Y(y_0,\beta)$.

PROOF Without loss of generality, we assume $x_0 = 0$ and $y_0 = 0$. Define

$$\Phi(x, y) = y - [D_y F(0, 0)]^{-1} F(x, y), \qquad (x, y) \in U.$$

By (a) Φ is continuous from U into Y. Since

$$[D_y \Phi(0, 0)]^{-1} \big(D_y \Phi(0, 0) - D_y \Phi(x, y)\big),$$

by (c) there is $\gamma > 0$ small enough such that

$$\|D_y \Phi(x, y)\|_{L(Y)} \le \frac{1}{2}, \qquad \forall\, (x, y) \in B_X(0, \gamma) \times B_Y(0, \gamma) \subset U.$$

Thus Lemma 2.2 and the continuity of Φ entail the inequality

$$\|\Phi(x, y_1) - \Phi(x, y_2)\|_Y \le \frac{1}{2}\|y_1 - y_2\|_Y, \qquad \|x\|_X, \|y_1\|_Y, \|y_2\|_Y \le \beta < \gamma.$$

Using now (a), we find $0 < \alpha < \beta$ such that

$$\|\Phi(x, 0)\|_Y \le \frac{\beta}{2}, \qquad \|x\|_X \le \alpha.$$

Then, for $\|x\|_X \le \alpha$ and $\|y\|_Y \le \beta$,

$$\|\Phi(x, y)\|_Y \le \|\Phi(x, 0)\|_Y + \|\Phi(x, y) - \Phi(x, 0)\|_Y \le \frac{1}{2}\big(\beta + \|y\|_Y\big) \le \beta.$$

Therefore the continuous map $\Phi : \overline{B}_X(0, \alpha) \times \overline{B}_Y(0, \beta) \to \overline{B}_Y(0, \beta)$ is a contraction on $\overline{B}_Y(0, \beta)$ uniformly in $\overline{B}_X(0, \alpha)$. From Corollary 1.4, there exists a unique continuous function $f : \overline{B}_X(0, \alpha) \to \overline{B}_Y(0, \beta)$ such that $\Phi(x, f(x)) = f(x)$, that is, $F(x, f(x)) = 0$. $\diamond$

Obviously, the thesis still holds replacing in the hypotheses closed balls with open balls.

Corollary *Let the hypotheses of Theorem 2.4 hold. If in addition F is Fréchet differentiable at $u_0 = (x_0, y_0)$, then f is Fréchet differentiable at x_0, and*

$$f'(x_0) = -[D_y F(u_0)]^{-1} D_x F(u_0).$$

PROOF Applying the definition of Fréchet differentiability to $F(x, f(x))$ at the point $(x_0, f(x_0))$, we get

$$0 = D_x F(u_0)(x - x_0) + D_y F(u_0)(f(x) - f(x_0)) + \sigma(x - x_0, f(x) - f(x_0)).$$

Notice that the above relation implies that f is locally Lipschitz at x_0. Hence

$$\frac{\|\sigma(x - x_0, f(x) - f(x_0))\|_Z}{\|x - x_0\|_X} \longrightarrow 0 \qquad \text{uniformly as} \qquad \|x - x_0\|_X \to 0$$

which yields the thesis. $\diamond$

A consequence of Theorem 2.4 is the inverse function theorem.

Theorem *Let X, Y be Banach spaces, $V \subset Y$ open, $y_0 \in V$. Let $g : V \to X$ be Fréchet differentiable in a neighborhood of y_0, $g(y_0) = x_0$, g' continuous at y_0, and $g'(y_0)$ invertible. Then there are $\alpha, \beta > 0$ and a unique continuous function $f : B_X(x_0, \alpha) \to B_Y(y_0, \beta)$ such that $x = g(f(x))$ for every $x \in B_X(x_0, \alpha)$. Moreover, f is Fréchet differentiable at x_0 and $f'(x_0) = g(y_0)^{-1}$.*

PROOF Apply Theorem 2.4 and the subsequent corollary to $F(x, y) = g(y) - x$, keeping in mind the considerations made in 2.3. $\diamond$

Theorem 2.4 can also be exploited to provide an alternative proof to the well-known fact that the set of invertible bounded linear operators between Banach spaces is open.

Theorem *Let X, Y be Banach spaces, and let $L_{\mathrm{reg}}(X, Y) \subset L(X, Y)$ be the set of invertible bounded linear operators from X onto Y. Then $L_{\mathrm{reg}}(X, Y)$ is open in $L(X, Y)$. Moreover, the map $T \mapsto T^{-1}$ is continuous.*

PROOF Let $F : L(X, Y) \times L(Y, X) \to L(X)$ defined by $F(T, S) = \mathbb{I}_Y - TS$. Let $T_0 \in L_{\mathrm{reg}}(X, Y)$, and set $S_0 = T_0^{-1}$. Notice that $D_S(T, S)(R) = -TR$. In particular, $D_S(T_0, S_0)(R) = -T_0 R$. Then the hypotheses of Theorem 2.4 are satisfied; therefore there is a continuous function $f : B_{L(X,Y)}(T_0, \alpha) \to L(Y, X)$ such that $\mathbb{I}_Y - Tf(T) = 0$, that is $Tf(T) = \mathbb{I}_Y$. Analogously, we can find a continuous function $f_1 : B_{L(X,Y)}(T_0, \alpha) \to L(Y, X)$ (perhaps for a smaller α) such that $f_1(T)T = \mathbb{I}_X$. It is straightforward to verify that $f \equiv f_1$, that is, $f(T) = T^{-1}$ for all $T \in B_{L(X,Y)}(T_0, \alpha)$. $\diamond$

2.5 Location of zeros Let X, Y be Banach spaces, and $f : B_X(x_0, r) \to Y$ be a Fréchet differentiable map. In order to find a zero for f, the idea is to apply an iterative method constructing a sequence x_n (starting from x_0) so that x_{n+1} is the zero of the tangent of f at x_n. Assuming that $f'(x)^{-1} \in L(Y, X)$ on $B_X(x_0, r)$, one has

$$x_{n+1} = x_n - f'(x_n)^{-1} f(x_n) \tag{1}$$

provided $x_n \in B_X(x_0, r)$ for every n. This procedure is known as the *Newton method*. However, for practical purposes, it might be complicated to invert f' at each step. So one can try the modification

$$x_{n+1} = x_n - f'(x_0)^{-1} f(x_n). \tag{2}$$

Clearly, using (2) in place of (1), a lower convergence rate is to be expected.

The following result is based on (2).

Theorem *Let X, Y be Banach spaces, and $f : B_X(x_0, r) \to Y$ be a Fréchet differentiable map. Assume that, for some $\lambda > 0$,*

(*a*) *$f'(x_0)$ is invertible;*

(*b*) $\|f'(x) - f'(x_0)\|_{L(X,Y)} \le \lambda \|x - x_0\|_X, \quad \forall\, x \in B_X(x_0, r)$;

(*c*) $\mu := 4\lambda \|f'(x_0)^{-1}\|^2_{L(Y,X)} \|f(x_0)\|_Y \le 1$;

(*d*) $s := 2\|f'(x_0)^{-1}\|_{L(Y,X)} \|f(x_0)\|_Y < r$.

Then there exists a unique $\bar{x} \in \overline{B}_X(x_0, s)$ *such that* $f(\bar{x}) = 0$.

PROOF Define $\Phi : \overline{B}_X(x_0, s) \to X$ as $\Phi(x) = x - f'(x_0)^{-1} f(x)$. Then

$$\|\Phi'(x)\|_{L(X)} \le \|f'(x_0)^{-1}\|_{L(Y,X)} \|f'(x_0) - f'(x)\|_{L(X,Y)} \le \lambda s \|f'(x_0)^{-1}\|_{L(Y,X)} = \frac{\mu}{2}.$$

Hence Φ is Lipschitz, with Lipschitz constant less than or equal to $\mu/2 \le 1/2$. Moreover,

$$\|\Phi(x_0) - x_0\|_X \le \|f'(x_0)^{-1}\|_{L(Y,X)} \|f(x_0)\|_Y = \frac{s}{2}$$

which in turn gives

$$\|\Phi(x) - x_0\|_X \le \|\Phi(x) - \Phi(x_0)\|_X + \|\Phi(x_0) - x_0\|_X \le \frac{\mu}{2}\|x - x_0\|_X + \frac{s}{2} \le s.$$

Hence Φ is a contraction on $\overline{B}_X(x_0, s)$. From Theorem 1.3 there exists a unique $\bar{x} \in \overline{B}_X(x_0, s)$ such that $\Phi(\bar{x}) = \bar{x}$, which implies $f(\bar{x}) = 0$. ◇

Concerning the convergence speed of x_n to $\bar{x}$, by virtue of the remark after Theorem 1.3, we get

$$\|x_n - \bar{x}\|_X \le \frac{s\mu^n}{(2 - \mu)2^n}.$$

Also, since

$$x_{n+1} - \bar{x} = f'(x_0)^{-1}(f'(x_0) - f'(x_n))(x_n - \bar{x}) + o(\|x_n - \bar{z}\|_X)$$

it follows that

$$\|x_{n+1} - \bar{x}\|_X = \frac{\mu}{2}\|x_n - \bar{x}\|_X + o(\|x_n - \bar{z}\|_X).$$

Hence

$$\|x_{n+1} - \bar{x}\|_X \le c\|x_n - \bar{x}\|_X$$

for some $c \in (0, 1)$. for all large n. This is usually referred to as *linear convergence* of the method.

Remark If we take $\mu < 2$, and we assume that f' is Lipschitz continuous on $B_X(x_0, r)$ with Lipschitz constant λ, we can still obtain the thesis with an entirely different proof (see, e.g., [2], pp.157–159), exploiting the iterative method (1). In this case we get the much better estimates

$$\|x_n - \bar{x}\|_X \le \frac{s}{2^n}\left(\frac{\mu}{2}\right)^{2^n - 1}$$

and

$$\|x_{n+1} - \bar{x}\|_X \le c\|x_n - \bar{x}\|^2_X$$

for some $c > 0$ (i.e., we have *quadratic convergence*).

Ordinary differential equations in Banach spaces

2.6 The Riemann integral Let X be a Banach space, $I = [\alpha, \beta] \subset \mathbb{R}$. The notion of Riemann integral and the related properties can be extended with no differences from the case of real-valued functions to X-valued functions on I. In particular, if $f \in C(I, X)$ then f is Riemann integrable on I,

$$\left\| \int_\alpha^\beta f(t)\, dt \right\|_X \leq \int_\alpha^\beta \|f(t)\|_X dt$$

and

$$\frac{d}{dt} \int_\alpha^t f(y)\, dy = f(t), \qquad \forall\, t \in I.$$

Recall that a function $h : I \to X$ is differentiable at $t_0 \in I$ if the limit

$$\lim_{t \to t_0} \frac{h(t) - h(t_0)}{t - t_0}$$

exists in X. This limit is the derivative of h at t_0, and is denoted by $h'(t_0)$ or $\frac{d}{dt}h(t_0)$. If $t_0 \in (\alpha, \beta)$ we recover the definition of Fréchet differentiability.

It is easy to see that if $h'(t) = 0$ for all $t \in [\alpha, \beta]$, then $h(t)$ is constant on $[\alpha, \beta]$. Indeed, for every $\Lambda \in X^*$, we have $(\Lambda \circ h)'(t) = \Lambda' h(t) = 0$, that implies $\Lambda(h(t) - h(\alpha)) = 0$, and from the Hahn-Banach theorem there is $\Lambda \in X^*$ such that $\Lambda(h(t) - h(\alpha)) = \|h(t) - h(\alpha)\|$.

2.7 The Cauchy problem Let X be a Banach space, $U \subset \mathbb{R} \times X$, U open, $u_0 = (t_0, x_0) \in U$, $f : U \to X$ continuous. The problem is to find a closed interval I, with t_0 belonging to the interior of I and a differentiable function $x : I \to X$ such that

$$\begin{cases} x'(t) = f(t, x(t)), & t \in I \\ x(t_0) = x_0. \end{cases} \tag{3}$$

It is apparent that such x is automatically of class C^1 on I. Also, it is readily seen that (3) is equivalent to the integral equation

$$x(t) = x_0 + \int_{t_0}^t f(y, x(y))\, dy, \qquad t \in I. \tag{4}$$

Namely, x is a solution to (3) if and only if it is a solution to (4).

2.8 Theorem [Local solution] *Assume the following hypotheses:*

(a) *f is continuous;*

(b) *The inequality*

$$\|f(t, x_1) - f(t, x_2)\|_X \leq k(t)\|x_1 - x_2\|_X, \qquad \forall\, (t, x_1), (t, x_2) \in U$$

holds for some $k(t) \in [0, \infty]$;

(*c*) $k \in L^1((t_0 - a, t_0 + a))$ *for some* $a > 0$;

(*d*) *There exist* $m \geq 0$ *and* $\overline{B}_{\mathbb{R}\times X}(u_0, s) \subset U$ *such that*

$$\|f(t,x)\|_X \leq m, \qquad \forall\, (t,x) \in \overline{B}_{\mathbb{R}\times X}(u_0, s).$$

Then there is $\tau_0 > 0$ *such that, for any* $\tau < \tau_0$, *there exists a unique solution* $x \in C^1(I_\tau, X)$ *to the Cauchy problem 2.7, with* $I_\tau = [t_0 - \tau, t_0 + \tau]$.

Remark Notice first that from (b), since U is open, k is defined in a neighborhood of zero. If k is constant then (c)-(d) are automatically satisfied. Indeed, for $(x,t) \in \overline{B}_{\mathbb{R}\times X}(u_0, s)$, we have

$$\|f(t,x)\|_X \leq ks + \max_{|t-t_0|\leq s} \|f(t,x_0)\|_X.$$

Also, (d) is always true if X is finite-dimensional, for closed balls are compact. In both cases, setting

$$s_0 = \sup\left\{\sigma > 0 : \overline{B}_{\mathbb{R}\times X}(u_0, \sigma) \subset U\right\}$$

we can choose any $s < s_0$.

PROOF Let $r = \min\{a, s\}$, and set

$$\tau_0 = \min\left\{r, \frac{r}{m}\right\}.$$

Select then $\tau < \tau_0$, and consider the complete metric space $Z = \overline{B}_{C(I_\tau,X)}(x_0, r)$ with the metric induced by the norm of $C(I_\tau, X)$ (here x_0 is the constant function equal to x_0). Since $\tau < r$, if $z \in Z$ then $(t, z(t)) \in \overline{B}_{\mathbb{R}\times X}(u_0, r) \subset U$ for all $t \in I_\tau$. Hence, for $z \in Z$, define

$$F(z)(t) = x_0 + \int_{t_0}^{t} f(y, z(y))\, dy, \qquad t \in I_\tau.$$

Observe that

$$\sup_{t\in I_\tau} \|F(z)(t) - x_0\|_X \leq \sup_{t\in I_\tau} \left| \int_{t_0}^{t} \|f(y, z(y))\|_X dy \right| \leq m\tau \leq r.$$

We conclude that F maps Z into Z. The last step is to show that F^n is a contraction on Z for some $n \in N$. By induction on n we show that, for every $t \in I_\tau$,

$$\|F^n(z_1)(t) - F^n(z_2)(t)\|_X \leq \frac{1}{n!}\left| \int_{t_0}^{t} k(y)\, dy \right|^n \|z_1 - z_2\|_{C(I_\tau,X)}. \tag{5}$$

For $n = 1$ it holds easily. So assume it is true for $n-1$, $n \geq 2$. Then, taking $t > t_0$ (the argument for $t < t_0$ is analogous),

$$\begin{aligned}
&\|F^n(z_1)(t) - F^n(z_2)(t)\|_X \\
&= \|F(F^{n-1}(z_1))(t) - F(F^{n-1}(z_2))(t)\|_X \\
&\leq \int_{t_0}^t \|f(y, F^{n-1}(z_1)(y)) - f(y, F^{n-1}(z_2)(y))\|_X dy \\
&\leq \int_{t_0}^t k(y)\|F^{n-1}(z_1)(y) - F^{n-1}(z_2)(y)\|_X dy \\
&\leq \frac{1}{(n-1)!}\left[\int_{t_0}^t k(y)\left(\int_{t_0}^y k(w)\,dw\right)^{n-1} dy\right]\|z_1 - z_2\|_{C(I_\tau,X)} \\
&= \frac{1}{n!}\left(\int_{t_0}^t k(y)\,dy\right)^n \|z_1 - z_2\|_{C(I_\tau,X)}.
\end{aligned}$$

Therefore from (5) we get

$$\|F^n(z_1) - F^n(z_2)\|_{C(I_\tau,X)} \leq \frac{1}{n!}\|k\|^n_{L^1(I_\tau)}\|z_1 - z_2\|_{C(I_\tau,X)}$$

which shows that for n big enough F^n is a contraction. By means of Corollary 1.5, we conclude that F admits a unique fixed point, which is clearly the (unique) solution to the integral equation (4) and hence to (3). ◇

2.9 Proposition [Continuous dependence] *The solution to the Cauchy problem 2.7 depends with continuity on the initial data.*

PROOF Assume that $x_j \in C(I_\tau, X)$ are two solutions to (3) with initial data x_{0j} $(j = 1, 2)$. Setting $x = x_1 - x_2$ and $x_0 = x_{01} - x_{02}$, from (b) we get

$$\|x(t)\|_X \leq \|x_0\|_X + \left|\int_{t_0}^t k(y)\|x(y)\|_X\,dy\right|, \qquad \forall\, t \in I_\tau. \tag{6}$$

The positive function

$$\varphi(t) = \|x_0\|_X \exp\left[\left|\int_{t_0}^t k(y)\,dy\right|\right]$$

satisfies the equation

$$\varphi(t) = \|x_0\|_X + \left|\int_{t_0}^t k(y)\varphi(y)\,dy\right|, \qquad \forall\, t \in I_\tau. \tag{7}$$

By comparing (6) with (7), we conclude that

$$\|x(t)\|_X \leq \|x_0\|_X \exp\left[\left|\int_{t_0}^t k(y)\,dy\right|\right], \qquad \forall\, t \in I_\tau. \tag{8}$$

Indeed, defining $\psi = \|x\|_X - \varphi$, addition of (6) and (7) entails

$$\psi(t) \leq \operatorname{sgn}(t - t_0) \int_{t_0}^{t} k(y)\psi(y)\,dy, \qquad \forall\, t \in I_\tau.$$

Let us show that the above inequality implies $\psi \leq 0$ on $[t_0, t_0 + \rho]$ for any $\rho < \tau$, an thus on $[t_0, t_0 + \tau)$ (the argument for $(t_0 - \tau]$ is the same). Choose n big enough such that

$$\int_{t_0 + \frac{j\rho}{n}}^{t_0 + \frac{(j+1)\rho}{n}} k(y)\;dy < 1, \qquad \forall\, j = 0, 1, \ldots, n - 1.$$

To finish the proof, we use an inductive argument. Suppose we proved that $\psi \leq 0$ on $[t_0, t_0 + j\rho/n]$ for some $j \leq n - 1$, and let t^* be such that

$$\psi(t^*) = \max\big\{\psi(t) : t \in [t_0 + j\rho/n, t_0 + (j+1)\rho/n]\big\}.$$

Then

$$\psi(t^*) \leq \int_{t_0}^{t^*} k(y)\psi(y)\,dy \leq \psi(t^*) \int_{t_0 + \frac{j\rho}{n}}^{t_0 + \frac{(j+1)\rho}{n}} k(y)\,dy.$$

So if $\psi(t^*)$ is strictly positive, we may cancel $\psi(t^*)$ in the above inequality, getting that the integral exceeds 1. Therefore $\psi \leq 0$ on $[t_0, t_0 + (j+1)\rho/n]$. ◇

Remark The implication (6)⇒(8) is known as the *Gronwall lemma*, which can also be proved via differential techniques.

2.10 Theorem [Global solution] *Let $U = (\alpha, \beta) \times X$. Assume (a) and (b) of Theorem 2.8, and replace (c) with*

(c′) $k \in L^1((\alpha, \beta))$.

Then there exists a unique solution $x \in C^1(I, X)$ to the Cauchy problem 2.7, for every $I \subset (\alpha, \beta)$.

PROOF Proceed like in the proof of Theorem 2.8, taking now $Z = C(I, X)$. The details are left to the reader. ◇

When f is merely continuous and fulfills a compactness property, it is possible to provide an existence result, exploiting the Schauder-Tychonoff fixed point theorem.

2.11 Theorem [Peano] *Assume f be a continuous function, and $f(V)$ be relatively compact in X for some open neighborhood $V \subset U$ of u_0. Then there exists $\tau > 0$ such that there is a (possibly nonunique) solution $x \in C^1(I_\tau, X)$ to the Cauchy problem 2.7, with $I_\tau = [t_0 - \tau, t_0 + \tau]$.*

PROOF Choose $r > 0$ such that $\overline{B}_{\mathbb{R}\times X}(u_0, r) \subset V$, and let

$$m = \sup\left\{\|f(t,x)\|_X : (x,t) \in \overline{B}_{\mathbb{R}\times X}(u_0, r)\right\}.$$

Finally, set $\tau = r/m$. With the same notation of the proof of Theorem 2.8, F maps the closed and convex set $Z = \overline{B}_{C(I_\tau, X)}(u_0, r)$ into itself. Moreover, F is continuous, since for $z_n, z \in Z$,

$$\sup_{t\in I_\tau} \|F(z_n)(t) - F(z)(t)\|_X \le \int_{t_0-\tau}^{t_0+\tau} \|f(y, z_n(y)) - f(y, z(y))\|_X dy$$

which vanishes as $z_n \to z$, due to the dominated convergence theorem. The Cauchy problem 2.7 has a solution if F has fixed point, which is guaranteed by Theorem 1.18 once we show that $F(Z)$ is relatively compact. To establish this fact, we appeal to the Ascoli theorem for X-valued continuous functions (see, e.g., Theorem 0.4.11 in [4]). First we observe that $F(Z)$ is an equicontinuous family, since for any $z \in Z$ and $t_1, t_2 \in I_\tau$,

$$\|F(z)(t_1) - F(z)(t_2)\|_X = \left\| \int_{t_1}^{t_2} f(y, z(y))\, dy \right\|_X \le m|t_1 - t_2|.$$

Next we prove that, for every fixed $t \in I_\tau$, $\bigcup_{z\in Z} F(z)(t)$ is relatively compact. From the hypotheses, there is a compact set $K \subset X$ such that $f(y, z(y)) \in K$ for all $y \in I_\tau$ and all $z \in Z$. Let $H = \mathrm{co}(K)$. It is well-known that in a Banach space the convex hull of a compact set is relatively compact (see, e.g., [14], Theorem 3.25). Then, for a fixed $t \in I_\tau$ and every $z \in Z$ we have

$$F(z)(t) = x_0 + \int_{t_0}^{t} f(y, z(y))\, dy \in x_0 + (t - t_0)\overline{H}$$

which is compact. Indeed, $\int_{t_0}^t f(y, z(y))\, dy$ is the limit of the Riemann sums $(t - t_0)\sum_i f(y_i, z(y_i))(y_i - y_{i-1})/(t - t_0)$. Hence Theorem 1.18 applies, and the thesis follows. $\diamond$

Definition The Peano theorem is said to hold in a Banach space X if for an arbitrary continuous function f and an arbitrary $x_0 \in X$ the Cauchy problem 2.7 admits a solution $x(t)$ in some neighborhood of t_0.

It is clear from Theorem 2.11 that if X is finite-dimensional the Peano theorem holds, since finite-dimensional Banach spaces have the Heine-Borel property. What about infinite-dimensional Banach spaces? Let us examine a famous counterexample to the existence problem.

2.12 Example [Dieudonné] Take the Banach space c_0 with the supremum norm and define the function $f : c_0 \to c_0$ by

$$(f(x))_n = \sqrt{|x_n|} + \frac{1}{1+n}, \qquad \text{for } x = (x_0, x_1, x_2, \ldots) \in c_0.$$

Consider the Cauchy problem

$$\begin{cases} x'(t) = f(x(t)), & t \in (-\varepsilon, \varepsilon) \\ x(0) = 0. \end{cases}$$

Due to uniform continuity of the real map $s \mapsto \sqrt{|s|}$, f is uniformly continuous on c_0. Assume now that the Cauchy problem admits a solution $y \in C^1((-\varepsilon, \varepsilon), c_0)$ for some $\varepsilon > 0$. In particular, each component y_n is differentiable in $(-\varepsilon, \varepsilon)$ and fulfills the Cauchy problem

$$\begin{cases} y_n'(t) = \sqrt{|y_n(t)|} + \dfrac{1}{1+n}, & t \in (-\varepsilon, \varepsilon) \\ y_n(0) = 0. \end{cases}$$

Then $y_n(t) > 0$ for $t \in (0, \varepsilon)$, and

$$2\sqrt{y_n(t)} - \frac{2}{1+n} \log\Big(\sqrt{y_n(t)} + \frac{1}{1+n}\Big) + \frac{2}{1+n} \log\Big(\frac{1}{1+n}\Big) = t, \qquad \forall\, t \in (0, \varepsilon).$$

On the other hand, $y(t) \in c_0$, that is, $\lim_{n\to\infty} y_n(t) = 0$. Therefore the limit as $n \to \infty$ of the left-hand side of the above equality must be 0 for all $t \in (0, \varepsilon)$. This is a contradiction, and we conclude that no such solution y exists.

In fact, it has been shown quite recently that the Peano theorem provides a characterization of finite-dimensional Banach spaces, namely,

2.13 Theorem [Godunov] *If the Peano theorem holds in a Banach space X, then X is finite-dimensional.*

Semilinear equations of evolution

2.14 Strongly continuous semigroups Let X be a Banach space. A one-parameter family $S(t)$ $(t \geq 0)$ of bounded linear operators on X is said to be a *strongly continuous semigroup* (*C_0-semigroup*, for short) if

(a) $S(0) = \mathbb{I}$ (identity operator on X);

(b) $S(t + s) = S(t)S(s)$ for every $t, s \geq 0$;

(c) $\lim_{t\to 0} S(t)x = x$ for every $x \in X$ (strong continuity).

As a quite direct application of the uniform boundedness theorem, there exist $\omega \geq 0$ and $M \geq 1$ such that

$$\|S(t)\|_{L(X)} \leq Me^{\omega t}, \qquad \forall\, t \geq 0.$$

This in turn entails the continuity of the map $t \mapsto S(t)x$ from $[0, \infty)$ to X, for every fixed $x \in X$ (cf. [9]).

The linear operator A of domain

$$\mathcal{D}(A) = \left\{ x \in X : \lim_{t\to 0} \frac{S(t)x - x}{t} \text{ exists} \right\}$$

defined by

$$Ax = \lim_{t\to 0} \frac{S(t)x - x}{t}, \qquad \forall\, x \in \mathcal{D}(A)$$

is the *infinitesimal generator* of the semigroup $S(t)$.

We now recall some basic facts on A (see, e.g., [9]).

Proposition *A is a closed linear operator with dense domain. For every fixed $x \in \mathcal{D}(A)$, the map $t \mapsto S(t)x$ belongs to $C^1([0,\infty), \mathcal{D}(A))$ and*

$$\frac{d}{dt} S(t)x = AS(t)x = S(t)Ax.$$

2.15 We consider the following semilinear Cauchy problem in X:

$$\begin{cases} x'(t) = Ax(t) + f(t, x(t)), & 0 < t \le T \\ x(0) = x_0 \in X. \end{cases} \tag{9}$$

where A is the infinitesimal generator of a strongly continuous semigroup $S(t)$, and $f : [0,T] \times X \to X$ is continuous and uniformly Lipschitz continuous on X with Lipschitz constant $\lambda \ge 0$.

Definition A function $x : [0,T] \to X$ is said to be a *classical solution* to (9) if it is differentiable on $[0,T]$, $x(t) \in \mathcal{D}(A)$ for every $t \in [0,T]$, and (9) is satisfied on $[0,T]$.

If x is a classical solution, it is necessarily unique, and it is given by

$$x(t) = S(t)x_0 + \int_0^t S(t-s) f(s, x(s))\, ds. \tag{10}$$

This can be easily proved integrating in ds on $[0,t]$ the derivative with respect to s of the (differentiable) function $S(t-s)x(s)$ and using (9). Notice that above (Riemann) integral is well-defined, since if $x \in C([0,T], X)$ the map $t \mapsto f(t, x(t))$ belong to $C([0,T], X)$ as well (see 2.6). Of course, there is no reason why there should exist a classical solution for a certain initial value x_0. However, formula (10) makes sense for every $x_0 \in X$. This motivates the following definition.

Definition A function $x : [0,T] \to X$ is said to be a *mild solution* to (9) if it continuous on $[0,T]$ and fulfills the integral equation (10).

Theorem *For every $x_0 \in X$ the Cauchy problem (9) has a unique mild solution. Moreover the map $x_0 \mapsto x(t)$ is Lipschitz continuous from X into $C([0,T], X)$.*

PROOF For a given $x_0 \in X$ define the map $F : C([0,T],X) \to C([0,T],X)$ by

$$F(x)(t) = S(t)x_0 + \int_0^t S(t-s)f(s,x(s))\,ds.$$

Then we have

$$\|F(x)(t) - F(y)(t)\|_X \le \lambda M t \|x-y\|_{C([0,T],X)}$$

where $M = \sup_{t\in[0,T]} \|S(t)\|_{L(X)}$. By an inductive argument, analogous to the one used in the proof of Theorem 2.8, we get

$$\|F^n(x)(t) - F^n(y)(t)\|_X \le \frac{(\lambda MT)^n}{n!}\|x-y\|_{C([0,T],X)}.$$

Hence for $n \in \mathbb{N}$ big enough F^n is a contraction, so by Corollary 1.5 F has a unique fixed point in $C([0,T],X)$ which is clearly the desired mild solution to the Cauchy problem (9).
To complete the proof, let y be the unique mild solution corresponding to the initial value y_0. Then

$$\|x(t) - y(t)\|_X \le M\|x_0 - y_0\|_X + \lambda M \int_0^t \|x(s) - y(s)\|_X\,ds.$$

By the Gronwall Lemma (cf. Proposition 2.9), we get at once

$$\|x(t) - y(t)\|_X \le M e^{\lambda MT}\|x_0 - y_0\|_X, \qquad \forall\, t \in [0,T]$$

which entails the Lipschitz continuity of the map $x_0 \mapsto x(t)$ ◇

An abstract elliptic problem

Let X, V be Banach spaces with compact and dense embeddings $V \hookrightarrow X \hookrightarrow V^*$. Assume we are given a bounded linear operator $A : V \to V^*$ and a (nonlinear) continuous map $B : X \to V^*$ which carries bounded sets into bounded sets, such that

$$\langle Au, u\rangle \ge \varepsilon\|u\|_V^2, \qquad \forall\, u \in V \tag{11}$$

and

$$\langle B(u), u\rangle \ge -c(1 + \|u\|_V^\alpha), \qquad \forall\, u \in X \tag{12}$$

for some $\varepsilon > 0$, $c \ge 0$ and $\alpha \in [0,2)$, where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between V^* and V.

Problem *Given $g \in V^*$, find a solution $u \in V$ to the abstract equation*

$$Au + B(u) = g. \tag{13}$$

For $v \in X$, let w be the solution to the equation

$$Aw = g - B(v).$$

From (11) A is injective onto V^*, and by the open mapping theorem, A^{-1} is a bounded linear operator from V^* onto V. Therefore

$$w = A^{-1}(f - B(v)) \in V.$$

Define then the map $f : X \to X$ as $f(v) = A^{-1}(g - B(v))$. Notice that f is continuous and compact. Suppose then that, for some λ, there is u_λ such that $u_\lambda = \lambda f(u_\lambda)$. This means that u_λ solves the equation

$$Au_\lambda + \lambda B(u_\lambda) = \lambda g.$$

Taking the duality pairing of the above equation and u_λ, and exploiting (11)-(12), we get

$$\varepsilon \|u_\lambda\|_V^2 \le \lambda c(1 + \|u_\lambda\|_V^\alpha) + \lambda \|g\|_{V^*} \|u_\lambda\|_V.$$

Recalling now that $\lambda \in [0, 1]$, and using the Young inequality

$$ab \le \nu a^p + K(\nu, p) b^q \qquad (a, b \ge 0,\ \nu > 0)$$

where $K(\nu, p) = (\nu p)^{-q/p} q^{-1}$ $(1 < p,\ q < \infty,\ 1/p + 1/q = 1)$, we find the *a priori* estimate

$$\|u_\lambda\|_V^2 \le \frac{2}{\varepsilon}\Big(c + K(\varepsilon/4, 2/\alpha)\, c^{2/(2-\alpha)} + \frac{1}{\varepsilon}\|g\|_{V^*}^2\Big).$$

A direct application of Theorem 1.20 entails the existence of a fixed point u for f, which is clearly a solution to (13). Notice that the solution might not be unique.

Example Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with smooth boundary $\partial\Omega$. Find a weak solution to the nonlinear elliptic problem

$$\begin{cases} -\Delta u + u^5 = g \\ u_{|\partial\Omega} = 0. \end{cases}$$

In this case $V = H_0^1(\Omega)$, $X = L^6(\Omega)$, and $g \in H_0^1(\Omega)^* = H^{-1}(\Omega)$ (see, e.g., [1, 5] for the definitions and the properties of the Sobolev spaces H_0^1 and H^{-1}; in particular, we recall that the embedding $H_0^1(\Omega) \hookrightarrow L^6(\Omega)$ is compact when $\Omega \subset \mathbb{R}^2$). Then $A = -\Delta$ and $B(u) = u^5$ are easily seen to fulfill the required hypotheses.

The invariant subspace problem

The invariant subspace problem is probably *the* problem of operator theory. The question, that attracted the attention of a great deal of mathematicians, is quite simple to state. Given a Banach space X and an operator $T \in L(X)$, find a closed nontrivial subspace M of X (i.e., $M \neq X$ and $M \neq \{0\}$) for which $TM \subset M$. Such

M is said to be an *invariant subspace* for T. It is known that not all continuous linear operators on Banach spaces have invariant subspaces. The question is still open for Hilbert spaces.

The most general, and at the same time most spectacular, result on the subject, is the Lomonosov theorem, that provides the existence of hyperinvariant subspaces for a vast class of operators. The proof is relatively simple, and the role played by the Schauder-Tychonoff theorem is essential. In order to state the result, we need first a definition.

Definition Let X be a Banach space. An invariant subspace M for $T \in L(X)$ is said to be *hyperinvariant* if it is invariant for all operators commuting with T (that is, for all $T' \in L(X)$ such that $TT' = T'T$).

Remark If $T \in L(H)$ is *nonscalar*, i.e., is not a multiple of the identity, and it has an eigenvalue λ, then the eigenspace M corresponding to λ is hyperinvariant for T. Indeed, if $x \in M$ and T' commutes with T, we have that

$$\lambda T'x = T'Tx = TT'x.$$

Therefore $T'x \in M$.

2.16 Theorem [Lomonosov] *Let X be a Banach space. Let $T \in L(X)$ be a nonscalar operator commuting with a nonzero compact operator $S \in L(X)$. Then T has a hyperinvariant subspace.*

We anticipate an observation that will be used in the proof.

Remark Let $S \in L(X)$ be a compact operator (cf. Definition 1.19). If $\lambda \neq 0$ is an eigenvalue of S, then the eigenspace

$$F := \{x \in X : Sx = \lambda x\}$$

relative to λ has finite dimension. Indeed, the restriction of S on F is a (nonzero) multiple of the identity on F, and the identity is compact if and only if the space is finite-dimensional.

PROOF We proceed by contradiction. Let $\mathcal{A}$ be the algebra of operators commuting with T. It is immediate to see that if T has no hyperinvariant subspaces, then $\overline{\mathcal{A}x} = X$ for every $x \in X$, $x \neq 0$.
Without loss of generality, let $\|S\|_{L(X)} \leq 1$. Choose then $x_0 \in X$ such that $\|Sx_0\| > 1$ (which implies $\|x_0\| > 1$) and set $B = \overline{B}_X(x_0, 1)$. For $x \in \overline{SB}$ (notice that x cannot be the zero vector), there is $T' \in \mathcal{A}$ such that $\|T'x - x_0\| < 1$. Hence every $x \in \overline{SB}$ has an open neighborhood V_x such that $T'V_x \subset B$ for some $T' \in \mathcal{A}$. Exploiting the compactness of $\overline{SB}$, we find a finite cover $V_1, \ldots, V_n$ and $T_1', \ldots, T_n' \in \mathcal{A}$ such that

$$T_j'V_j \subset B, \qquad \forall\, j = 1, \ldots, n.$$

Let $\varphi_1, \ldots, \varphi_n \in C(\overline{SB})$ be a partition of the unity for $\overline{SB}$ subordinate to the open cover $\{V_j\}$, and define, for $x \in B$,

$$f(x) = \sum_{j=1}^{n} \varphi_j(Sx) T_j' Sx.$$

Then f is a continuous function from B into B. Since $T_j'S$ is a compact map for every j, it is easily seen that $f(B)$ is relatively compact. Hence Theorem 1.18, yielding the existence of $\bar{x} \in B$ such that $f(\bar{x}) = \bar{x}$. Defining the operator $\widetilde{T} \in \mathcal{A}$ as

$$\widetilde{T} = \sum_{j=1}^{n} \varphi_j(S\bar{x}) T_j'$$

we get the relation

$$\widetilde{T} S \bar{x} = \bar{x}.$$

But $\widetilde{T}S$ is a compact operator, hence the eigenspace F of $\widetilde{T}S$ relative to the eigenvalue 1 is finite-dimensional. Since $\widetilde{T}S$ commutes with T, we conclude that F is invariant for T, which means that T has an eigenvalue, and thus a hyperinvariant subspace, contrary to our assumption. ◇

Measure preserving maps on compact Hausdorff spaces

Let X be a compact Hausdorff space, and let $P(X)$ be the set of all Borel probability measures on X. By means of the Riesz representation theorem, the dual space of $C(X)$ can be identified with the space $M(X)$ of complex regular Borel measures on X. Recall that the norm $\|\mu\|$ of an element $\mu \in M(X)$ is given by the total variation of μ. It is straightforward to check that $P(X)$ is convex and closed in the weak* topology. Moreover, $P(X)$ is weak* compact. Indeed, it is a weak* closed subset of the unit ball of $M(X)$, which is weak* compact by the Banach-Alaoglu theorem.

Definition Let $\mu \in P(X)$. A μ-measurable map $f : X \to X$ is said to be *measure preserving* with respect to μ if $\mu(B) = \mu(f^{-1}(B))$ for every Borel set $B \subset X$. Such μ is said to be an *invariant measure* for f.

Notice that, if f is a μ-measurable map, the measure $\tilde{f}\mu$ defined by

$$\tilde{f}\mu(B) = \mu(f^{-1}(B)), \qquad \forall \text{ Borel set } B$$

belongs to $P(X)$. In particular, if f is continuous (and therefore measurable with respect to every $\mu \in P(X)$), we have a map $\tilde{f} : P(X) \to P(X)$ defined by $\mu \mapsto \tilde{f}\mu$. In addition, applying the monotone convergence theorem to an increasing sequence of simple functions, it is easy to see that

$$\int_X g \, d(\tilde{f}\mu) = \int_X g \circ f \, d\mu, \qquad \forall \, g \in C(X).$$

Lemma *Assume $f : X \to X$ be continuous. Then the map $\tilde{f} : P(X) \to P(X)$ is continuous in the weak* topology.*

PROOF Let $\{\mu_\iota\}_{\iota \in I} \subset P(X)$ be a net converging to some $\mu \in P(X)$. Then, for every $g \in C(X)$,

$$\lim_{\iota \in I} \int_X g \, d(\tilde{f}\mu_\iota) = \lim_{\iota \in I} \int_X g \circ f \, d\mu_\iota = \int_X g \circ f \, d\mu = \int_X g \, d(\tilde{f}\mu)$$

which entails the claimed continuity. ◇

We are now interested to find elements of $P(X)$ that are invariant measures for f. This is the same as finding a fixed point for the map $\tilde{f}$.

Theorem *Let $f : X \to X$ be a continuous map. Then there exists $\mu \in P(X)$ for which f is measure preserving.*

PROOF On account of the above discussion, $\tilde{f}$ is a continuous map of a compact convex subset of $M(X)$ into itself, and the existence of a fixed point is then guaranteed by Theorem 1.18. ◇

Clearly, the invariant measure for f might not be unique. In particular, notice that if x is a fixed point for f, then the Dirac measure δ_x is invariant for f.

Invariant means on semigroups

Let S be a semigroup, that is, a set endowed with an associative binary operation, and consider the (real) Banach space of all real-valued bounded functions on S, namely,

$$\ell^\infty(S) = \Big\{ f : S \to \mathbb{R} : \|f\| := \sup_{s \in S} |f(s)| < \infty \Big\}.$$

An element $f \in \ell^\infty(S)$ is positive if $f(s) \geq 0$ for every $s \in S$. A linear functional $\Lambda : \ell^\infty(S) \to \mathbb{R}$ is positive if $\Lambda f \geq 0$ for every positive element $f \in \ell^\infty(S)$. We agree to denote a constant function on S by the value of the constant.

We now recall a result that actually holds for more general situations.

Lemma *Let $\Lambda \in \ell^\infty(S)^*$, with $\|\Lambda\| = \Lambda 1 = 1$. Then Λ is positive.*

PROOF Assume not. Then there is $f \in \ell^\infty(S)$, $f \geq 0$, such that $\Lambda f = \beta < 0$. For $\varepsilon > 0$ small, we have

$$\|1 - \varepsilon f\| = \sup_{s \in S} |1 - \varepsilon f(s)| \leq 1.$$

Hence

$$1 < 1 - \varepsilon\beta \leq |1 - \varepsilon\beta| = |\Lambda(1 - \varepsilon f)| \leq \|1 - \varepsilon f\| \leq 1$$

leading to a contradiction. ◇

If $t \in S$, we can define the *left t-translation operator* $L_t : \ell^\infty(S) \to \ell^\infty(S)$ to be

$$(L_t f)(s) = f(ts), \qquad \forall\, s \in S.$$

In an analogous manner, we can define the *right t-translation operator* R_t.

Definition A (left) *invariant mean* on S is a positive linear functional Λ on $\ell^\infty(S)$ satisfying the following conditions:

(a) $\Lambda 1 = 1$;

(b) $\Lambda(L_s f) = \Lambda f$ for every $s \in S$ and every $f \in \ell^\infty(S)$.

When such a functional exists, S is said to be (left) *amenable*.

Clearly, we can give the above definition replacing left with right or two-sided. The distinction is relevant if S is not abelian.

2.17 Example [Banach] Let $S = \mathbb{N}$. Then $\ell^\infty(\mathbb{N}) = \ell^\infty$. An invariant mean in this case is called a *Banach generalized limit.* The reason is that if Λ is an invariant mean on $\mathbb{N}$ and $x = \{x_n\}_{n\in\mathbb{N}} \in \ell^\infty$ is such that $\lim_{n\to\infty} x_n = \alpha \in \mathbb{R}$, then $\Lambda x = \alpha$. Indeed, for any $\varepsilon > 0$, we can choose n_0 such that $\alpha + \varepsilon \le x_n \le \alpha + \varepsilon$ for every $n \ge n_0$. Hence, if we define $y = \{y_n\}_{n\in\mathbb{N}} \in \ell^\infty$ by $y_n = x_{n+n_0}$, we have $\Lambda x = \Lambda y$, and

$$\alpha - \varepsilon = \Lambda(\alpha - \varepsilon) \le \Lambda y \le \Lambda(\alpha + \varepsilon) \le \alpha + \varepsilon$$

which yields the equality $\Lambda x = \alpha$. To prove the existence of an invariant mean, one has to consider the subspace $\mathcal{M}$ of ℓ^∞ given by

$$\mathcal{M} = \left\{ x = \{x_n\}_{n\in\mathbb{N}} \in \ell^\infty : \lim_{n\to\infty} \frac{x_0 + \cdots + x_n}{n+1} = \alpha_x \in \mathbb{R} \right\}$$

and define the linear functional Λ_0 on $\mathcal{M}$ as $\Lambda_0 x = \alpha_x$. Setting for every $x \in \ell^\infty$

$$p(x) = \limsup_{n\to\infty} \frac{x_0 + \cdots + x_n}{n+1}$$

it is then possible, by means of the Hahn-Banach theorem, to extend Λ_0 to a functional Λ defined on the whole space, in such a way that $-p(-x) \le \Lambda x \le p(x)$ for every $x \in \ell^\infty$. In particular, Λ is continuous. A remarkable consequence of this fact is that not every continuous linear functional on ℓ^∞ can be given the representation

$$\{x_n\}_{n\in\mathbb{N}} \mapsto \sum_{n=0}^{\infty} c_n x_n$$

for some numerical sequence c_n. Indeed, for every $k \in \mathbb{N}$, taking $e_k = \{\delta_{nk}\}_{n\in\mathbb{N}}$, we have $\Lambda e_k = 0$. Hence if Λ has the above representation, all the c_n must be zero, i.e., Λ is the null functional, contrary to the fact that $\Lambda 1 = 1$.

The next result, based on the Markov-Kakutani theorem, provides an elegant generalization of Example 2.17.

2.18 Theorem [Day] *Let S be an abelian semigroup. Then S is amenable.*

PROOF Denote

$$K = \{\Lambda \in \ell^\infty(S)^* : \|\Lambda\| = \Lambda 1 = 1\}.$$

In particular, if $\Lambda \in K$ then Λ is positive. K is convex, and from the Banach-Alaoglu theorem is compact in the weak* topology of $\ell^\infty(S)^*$. We define the family of linear operators $T_s : \ell^\infty(S)^* \longrightarrow \ell^\infty(S)^*$, $s \in S$, as

$$(T_s\Lambda)(f) = \Lambda(L_s f), \qquad \forall\, f \in \ell^\infty(S).$$

First we show that T_s is continuous in the weak* topology for every $s \in S$. Of course, it is enough to show the continuity at zero. Thus let V be a neighborhood of zero of the local base for the weak* topology, that is,

$$V = \{\Lambda \in \ell^\infty(S)^* : |\Lambda f_j| < \varepsilon_j,\ j = 1, \ldots, n\}$$

for some $\varepsilon_1, \ldots, \varepsilon_n > 0$ and $f_1, \ldots, f_n \in \ell^\infty(S)$. Then

$$\begin{aligned} T_s^{-1}(V) &= \{\Lambda \in \ell^\infty(S)^* : |(T_s\Lambda)(f_j)| < \varepsilon_j,\ j = 1, \ldots, n\} \\ &= \{\Lambda \in \ell^\infty(S)^* : |\Lambda(L_s f_j)| < \varepsilon_j,\ j = 1, \ldots, n\} \end{aligned}$$

is an open neighborhood of zero as well.
The second step is to prove that $T_s K \subset K$. Indeed,

$$(T_s\Lambda)(1) = \Lambda(L_s 1) = \Lambda 1 = 1$$

and

$$\|T_s\Lambda\| = \sup_{\|f\|\leq 1} |(T_s\Lambda)(f)| = \sup_{\|f\|\leq 1} |\Lambda(L_s f)| \leq \sup_{\|f\|\leq 1} |\Lambda f| = \|\Lambda\| = 1$$

since $\|L_s f\| \leq \|f\|$.
Finally, for every $s, t \in S$,

$$T_s T_t \Lambda = T_s(\Lambda \circ L_t) = \Lambda \circ L_t \circ L_s = \Lambda \circ L_{st} = \Lambda \circ L_{ts} = T_t T_s \Lambda.$$

Hence by Theorem 1.23 there is $\Lambda \in K$ such that $T_s\Lambda = \Lambda$ for every $s \in S$, which means that $\Lambda(L_s f) = \Lambda f$ for every $s \in S$ and every $f \in \ell^\infty(S)$. ◇

Haar measures

2.19 Topological groups A *topological group* is a group G endowed with a Hausdorff topology that makes the group operations continuous; namely, the map $(x, t) \mapsto xy^{-1}$ is continuous for every $x, y \in G$.

For any $y \in G$ the maps $x \mapsto xy$, $x \mapsto yx$, and $x \mapsto x^{-1}$ are homeomorphisms of G onto G. Hence the topology of G is uniquely determined by any local base at the identity element e. Indeed, if U is a neighborhood of some $x \in G$, the sets $xU = \{xy : y \in U\}$ and $Ux = \{yx : y \in U\}$ are neighborhoods of e.

The next topological lemma will be used in the proof of the main result of the section.

Lemma *Let G be a compact topological group, and let $K_1, K_2 \subset G$ be disjoint compact sets. Then there exists an open neighborhood U of e such that no translate of U meets both K_1 and K_2.*

PROOF Let $O \subset G$ be an open set such that $O \supset K_1$ and $O \cap K_2 = \emptyset$. For every $x \in K_1$, $x^{-1}O \in \mathcal{U}$. Using the continuity of the group operations, select $U_x \in \mathcal{U}$ such that $U_x = U_x^{-1}$ and $U_x U_x U_x \subset x^{-1}O$, and let $V_x = xU_x \cap U_x x$. Since $V_x \in \mathcal{U}$, it follows that $\bigcup_{x \in K_1} V_x$ is an open cover of K_1, and by compactness there are $x_1, \dots, x_n \in K_1$ such that $K_1 \subset \bigcup_{i=1}^n V_{x_i}$. Finally define

$$U = \bigcap_{i=1}^n V_{x_i} \in \mathcal{U}.$$

Assume now that some translate of U meets both K_1 and K_2. Then there are $k_1 \in K_1$ and $k_2 \in K_2$ such that either $k_2 = k_1 u^{-1} v$ or $u^{-1} v k_2 = k_1$ for some $u, v \in U$. Let us examine the first case (being the second analogous). We have that $k_1 \in V_{x_i} \subset x_i U_{x_i}$ for some $i \in \{1, \dots, n\}$. Hence

$$k_2 \in x_i U_{x_i} U^{-1} U \subset x_i U_{x_i} (x_i U_{x_i})^{-1} x_i U_{x_i} = x_i U_{x_i} U_{x_i} U_{x_i} \subset x_i x_i^{-1} O \subset O$$

which is a contradiction. $\diamond$

Notice that the proof holds in fact for locally compact topological groups.

2.20 Definition Let G be a compact topological group. A *Haar measure* on G is a regular Borel probability measure μ which is simultaneously *left invariant*, i.e.,

$$\int_G f(x)\, d\mu(x) = \int_G f(yx)\, d\mu(x), \qquad \forall\, y \in G,\ f \in C(G) \tag{14}$$

and *right invariant*, i.e.,

$$\int_G f(x)\, d\mu(x) = \int_G f(xy)\, d\mu(x), \qquad \forall\, y \in G,\ f \in C(G) \tag{15}$$

and satisfies the relation

$$\int_G f(x)\, d\mu(x) = \int_G f(x^{-1})\, d\mu(x), \qquad \forall\, f \in C(G). \tag{16}$$

It is readily seen that a Haar measure on G, if it exists, is unique. Indeed, if μ and ν are two Borel probability measures on G, with μ left invariant and ν right invariant, for every $f \in C(G)$ the Fubini theorem yields

$$\begin{aligned}\int_G f(x)\,d\mu(x) &= \int_G f(yx)\,d\mu(x)\\ &= \int_G \left[\int_G f(yx)\,d\mu(x)\right] d\nu(y)\\ &= \int_G \left[\int_G f(yx)\,d\nu(y)\right] d\mu(x)\\ &= \int_G f(yx)\,d\nu(y)\\ &= \int_G f(y)\,d\nu(y)\end{aligned}$$

forcing the equality $\mu = \nu$.

2.21 Theorem *Let G be a compact topological group. Then there exists a unique Haar measure on G.*

PROOF By means of the above uniqueness argument, it is enough to prove the existence of two regular Borel probability measure μ and ν satisfying (14) and (15), respectively. Then (16) follows at once for if μ is a Haar measure on G, so is ν defined by $d\nu(x) = d\mu(x^{-1})$.

We then proceed by proving the existence of μ (the proof for ν being the same). Let $\mathcal{K}$ be the family of compact subsets of G, and $\mathcal{U}$ be the family of open subsets of G containing e. If $K \in \mathcal{K}$ and $U \in \mathcal{U}$, we can always cover K by a finite number of translates of U. We define the *covering number* $[K : U]$ of K by U the smallest number of translates of U required to cover K. For every $K \in \mathcal{K}$, we introduce the normalized covering ratio

$$\xi_K(U) = \frac{[K:U]}{[G:U]}, \qquad \forall\, U \in \mathcal{U}.$$

It is immediate to verify that $0 \le \xi_K(U) \le 1$, $\xi_{K_1}(U) \le \xi_{K_2}(U)$ for $K_1 \subset K_2$, and $\xi_{xK}(U) = \xi_K(U)$ for all $x \in G$. Moreover, if $K_1, K_2 \in \mathcal{K}$ are disjoint sets, exploiting the lemma we find $U \in \mathcal{U}$ such that no translate of U meets both K_1 and K_2. Hence any covering of $K_1 \cup K_2$ by translates of U is the disjoint union of coverings of K_1 and K_2. The same holds replacing U with any $V \in \mathcal{U}$, $V \subset U$. Thus

$$\xi_{K_1 \cup K_2}(V) = \xi_{K_1}(V) + \xi_{K_2}(V), \qquad \forall\, V \in \mathcal{U},\, V \subset U. \tag{17}$$

Notice that $(\mathcal{U}, \cap)$ is an abelian semigroup, so Theorem 2.18 applies, yielding the existence of an invariant mean Λ on $\mathcal{U}$. Since $\xi_K : \mathcal{U} \to \mathbb{R}$ belongs to $\ell^\infty(\mathcal{U})$, we can define

$$\psi(K) = \Lambda \xi_K, \qquad \forall\, K \in \mathcal{K}.$$

The following hold:

$$\psi(\emptyset) = 0, \quad \psi(G) = 1; \tag{18}$$
$$\psi(K_1) \leq \psi(K_2), \qquad \text{if } K_1 \subset K_2; \tag{19}$$
$$\psi(xK) = \psi(K), \qquad \forall\, x \in G. \tag{20}$$
$$\psi(K_1 \cup K_2) = \psi(K_1) + \psi(K_2), \qquad \text{if } K_1 \cap K_2 = \emptyset. \tag{21}$$

Properties (18)-(20) are direct consequences of the definition of $\xi_K(U)$, whereas (21) follows from (17) and the fact that if $\xi \in \ell^\infty(\mathcal{U})$ and there is $U \subset \mathcal{U}$ such that $\xi(V) = 0$ for every $V \subset \mathcal{U}$, $V \subset U$, then $\Lambda\xi = 0$. Indeed, $\Lambda(R_U\xi) = \Lambda\xi$, and $(R_U\xi)(W) = \xi(U \cap W) = 0$ for every $W \subset \mathcal{U}$.

For each open set $O \subset G$ define

$$\mu^*(O) = \sup\{\psi(K) : K \subset O,\ K \text{ compact}\}$$

and for every $E \subset G$ let

$$\mu^*(E) = \inf\{\mu^*(O) : O \supset E,\ O \text{ open}\}.$$

By (18)-(20), $\mu^*(\emptyset) = 0$, $\mu^*(G) = 1$, $\mu^*(E_1) \leq \mu^*(E_2)$ if $E_1 \subset E_2 \subset G$, and μ^* is left invariant. Moreover μ^* is countably subadditive, i.e.,

$$\mu^*\Big(\bigcup_{j=1}^{\infty} E_j\Big) \leq \sum_{j=1}^{\infty} \mu^*(E_j). \tag{22}$$

To see that, let $\varepsilon > 0$, and choose open sets $O_j \subset E_j$ with $\mu^*(O_j) \leq \mu^*(E_j) + \varepsilon/2^j$, and a compact set $K \subset \bigcup_{j=1}^{\infty} O_j$. Due to compactness, $K \subset \bigcup_{j=1}^{n} O_j$ for some n. Consider a partition of the unity $\varphi_1, \ldots, \varphi_n$ for K subordinate to the open cover $\{O_1, \ldots, O_n\}$, and set $K_j = \{x \in G : \varphi_j(x) > 1/n\}$. The K_j are compact, $\bigcup_{j=1}^{n} K_j \supset K$, and $K_j \cap K \subset O_j$. Thus

$$\psi(K) \leq \sum_{j=1}^{n} \psi(K_j \cap K) \leq \sum_{j=1}^{n} \mu^*(O_j) \leq \sum_{j=1}^{\infty} \mu^*(E_j) + \varepsilon.$$

Taking the supremum over all such K, and letting $\varepsilon \to 0$, we get (22). We conclude that μ^* is an outer measure. Applying then the Carathéodory extension process, we build a measure μ, which coincides with μ^* on the measurable sets, namely, those sets $E \subset G$ such that

$$\mu^*(T) = \mu^*(T \cap E) + \mu^*(T \cap E^C), \qquad \forall\, T \subset G.$$

The proof is finished if we show that the open sets are measurable, since μ is then a Borel outer regular measure (by construction) and hence regular (being finite). So let O be an open set, and T be any set. Let $\varepsilon > 0$, and take an open set $A \supset T$ such that $\mu^*(A) \leq \mu^*(T) + \varepsilon$. Choose a compact set $K \subset A \cap O$ with

$\mu^*(A \cap O) \leq \psi(K) + \varepsilon$, and a compact set $K_0 \subset A \cap K^C$ with $\mu^*(A \cap K^C) \leq \psi(K_0) + \varepsilon$. Then

$$\begin{aligned}&\mu^*(T \cap O) + \mu^*(T \cap O^C) \leq \mu^*(A \cap O) + \mu^*(A \cap K^C) \leq \psi(K) + \psi(K_0) + 2\varepsilon \\ &= \psi(K \cup K_0) + 2\varepsilon \leq \mu^*(A) + 2\varepsilon \leq \mu^*(T) + 3\varepsilon.\end{aligned}$$

Since ε is arbitrary, the desired conclusion is immediate. ◇

A different proof of Theorem 2.21, that directly applies the Markov-Kakutani theorem, can be found in [14]. The advantage of the approach presented here is that it can be easily modified to extend the result to the locally compact case.

2.22 When G is a locally compact topological group it is still possible to talk of Haar measures. In this case we shall distinguish between left and right invariant ones.

Definition Let G be a locally compact topological group. A *left Haar measure* on G is a nonnull Radon measure μ on G satisfying (14). Analogously, a *right Haar measure* on G is a nonnull Radon measure ν on G satisfying (15).

Recall that a *Radon measure* is a positive linear functional on $C_c(G)$ (the space of compactly supported continuous functions on G). From the Riesz representation theorem every such functional can be uniquely represented by a positive outer regular Borel measure, finite on compact sets, for which open sets are inner regular (cf. [12]).

Theorem *Let G be a locally compact topological group. Then there exists left Haar measures and right Haar measures on G. Moreover, any two left (right) Haar measure differ by a multiplicative positive constant.*

The proof of the above theorem can be found in [11]. We list some important properties of Haar measures, that generalize (14)-(16) (cf. [4, 11]).

Theorem *Let G be a locally compact topological group.*

(a) *There exists a continuous strictly positive function Δ on G, with $\Delta(e) = 1$ and $\Delta(xy) = \Delta(x)\Delta(y)$ for all $x, y \in G$, such that if μ is any left Haar measure on G, then*

$$\int_G f(xy)\, d\mu(x) = \Delta(y) \int_G f(x)\, d\mu(x), \qquad \forall\, y \in G,\ f \in C_c(G).$$

(b) *For any left Haar measure μ on G there holds*

$$\int_G f(x^{-1})\, d\mu(x) = \int_G f(x)\, \Delta(x) d\mu(x), \qquad \forall\, f \in C_c(G).$$

(c) *For any left Haar measure μ and any right Haar measure ν on G there is $c > 0$ such that*

$$\int_G f(x)\, d\nu(x) = c \int_G f(x)\Delta(x)\, d\mu(x), \qquad \forall\, f \in C_c(G).$$

(*d*) *Given any any left or right Haar measure on G, every nonvoid open set has nonnull measure, and G has finite measure if and only if it is compact.*

Notice from (a)-(c) it follows that if ν is any right Haar measure on G then

$$\int_G f(yx)\, d\nu(x) = \frac{1}{\Delta(y)} \int_G f(x)\, d\nu(x), \qquad \forall\, y \in G,\ f \in C_c(G).$$

Remark The function Δ is termed the *modular function.* A locally compact topological group is called *unimodular* if there exists a bi-invariant Radon measure. This happens if and only if $\Delta \equiv 0$. Besides compact groups, abelian groups are clearly unimodular.

Game theory

We consider a game with $n \geq 2$ players, under the assumption that the players do not cooperate among themselves. Each players pursue a *strategy*, in dependence of the strategies of the other players. Denote the set of all possible strategies of the k^{th} player by K_k, and set $K = K_1 \times \cdots \times K_n$. An element $x \in K$ is called a *strategy profile.* For each k, let $f_k : K \to \mathbb{R}$ be the *loss function* of the k^{th} player. If

$$\sum_{k=1}^{n} f_k(x) = 0, \qquad \forall\, x \in K \tag{23}$$

the game is said to be of *zero-sum.* The aim of each player is to minimize his loss, or, equivalently, to maximize his gain.

Definition A *Nash equilibrium* is a strategy profile with the property that no player can benefit by changing his strategy, while all other players keep their strategies unchanged. In formulas, it is an element $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n) \in K$ such that

$$f_k(\bar{x}) \leq f_k(\bar{x}_1, \ldots, \bar{x}_{k-1}, x_k, \bar{x}_{k+1}, \ldots, \bar{x}_n), \qquad \forall\, x_k \in K_k \tag{24}$$

for every $k = 1, \ldots, n$.

Strictly speaking, a Nash equilibrium suggests a convenient "cautious" strategy to be adopted by each player in the game. We said *a* strategy rather than *the* strategy, since a Nash equilibrium (if it exists) might not be unique.

We need of course further hypotheses on the sets K_k and on the maps f_k. It is reasonable to assume that, with all the other strategies fixed, the loss function f_k has a small variation in correspondence of a small variation of x_k. Also, loosely speaking, it is assumed that the average of losses corresponding to two different strategies of the k^{th} player is grater than the loss corresponding to the "average" strategy. Convexity can suitably translate this issue.

The fundamental result of game theory is the following.

2.23 Theorem [Nash] *For every $k = 1, \dots, n$, let K_k be a nonvoid, compact and convex subset of a locally convex space X_k. Assume that, for every k, the loss function f_k is continuous on K. In addition, for every fixed $x_j \in K_j$ with $j \neq k$, the map*

$$f_k(x_1, \dots, x_{k-1}, \cdot, x_{k+1}, \dots, x_n) : K_k \to \mathbb{R}$$

is convex. Then there exists $\bar{x} \in K$ satisfying (24), i.e., there is a Nash equilibrium.

PROOF Define $\Phi : K \times K \to \mathbb{R}$ as

$$\Phi(x, y) = \sum_{k=1}^{n} \big[f_k(x) - f_k(x_1, \dots, x_{k-1}, y_k, x_{k+1}, \dots, x_n) \big].$$

Then Φ is continuous, and $\Phi(x, \cdot)$ is concave for every fixed $x \in K$. From Theorem 1.24 there exists $\bar{x} \in K$ such that

$$\sup_{y \in K} \Phi(\bar{x}, y) \leq \sup_{y \in K} \Phi(y, y) = 0.$$

In particular, if we set $\bar{y} = (\bar{x}_1, \dots, \bar{x}_{k-1}, x_k, \bar{x}_{k+1}, \dots, \bar{x}_n)$, for $x_k \in K_k$, we get

$$\Phi(\bar{x}, \bar{y}) \leq 0, \qquad \forall\, x_k \in K_k$$

which is nothing but (24). ◇

The hypotheses can be weakened if we consider a two-player zero-sum game (sometimes called a *duel*). In this case, on account of (23), we have

$$\Psi(x_1, x_2) := f_1(x_1, x_2) = -f_2(x_2, x_2).$$

Therefore we can repeat the above proof taking Ψ to be convex and lower semicontinuous in the first variable, and concave and upper semicontinuous in the second. Now Ψ is the loss function of the first player or, equivalently, the gain function of the second one.

2.24 Theorem [von Neumann] *Let $K_1 \subset X_1$ and $K_2 \subset X_2$ be as in Theorem 2.23. Let $\Psi : K_1 \times K_2 \to \mathbb{R}$ be such that*

(*a*) $\Psi(\cdot, x_2)$ *is lower semicontinuous and convex* $\forall\, x_2 \in K_2$;

(*b*) $\Psi(x_1, \cdot)$ *is upper semicontinuous and concave* $\forall\, x_1 \in K_1$.

Then there exists a Nash equilibrium $(\bar{x}_1, \bar{x}_2) \in K_1 \times K_2$.

PROOF In this case $\Phi : (K_1 \times K_2) \times (K_1 \times K_2) \to \mathbb{R}$ has the form

$$\Phi((x_1, x_2), (y_1, y_2)) = -\Psi(y_1, x_2) + \Psi(x_1, y_2)$$

and the argument of the proof of Theorem 2.23 applies. ◇

Theorem 2.24 is known in the literature as the *minimax* theorem. The reason is clear from the next corollary.

Corollary *In the hypotheses of Theorem 2.24, the equality*

$$\inf_{x_1 \in K_1} \sup_{x_2 \in K_2} \Psi(x_1, x_2) = \Psi(\bar{x}_1, \bar{x}_2) = \sup_{x_2 \in K_2} \inf_{x_1 \in K_1} \Psi(x_1, x_2)$$

holds true.

PROOF Define $g(x_1) = \sup_{x_2 \in K_2} \Psi(x_1, x_2)$ and $h(x_2) = \inf_{x_1 \in K_1} \Psi(x_1, x_2)$. Then for all $x_1 \in K_1$ and $x_2 \in K_2$ we have

$$h(x_2) \leq \Psi(x_1, x_2) \leq g(x_1)$$

which entails

$$\sup_{x_2 \in K_2} h(x_2) \leq \inf_{x_1 \in K_1} g(x_1).$$

On the other hand, by Theorem 2.24,

$$h(\bar{x}_2) = \inf_{x_1 \in K_1} \Psi(x_1, \bar{x}_2) = \Psi(\bar{x}_1, \bar{x}_2) = \sup_{x_2 \in K_2} \Psi(\bar{x}_1, x_2) = g(\bar{x}_1).$$

Hence

$$\sup_{x_2 \in K_2} h(x_2) \geq h(\bar{x}_2) = g(\bar{x}_1) \geq \inf_{x_1 \in K_1} g(x_1) \geq \sup_{x_2 \in K_2} h(x_2)$$

so that all the above inequalities are in fact equalities. $\diamond$

We conclude the section considering a duel game where the sets K_1 and K_2 of all possible strategies of each player are finite. We also assume that, for $k = 1, 2$, player k plays randomly the strategy $x_k \in K_k$ with probability $p_k(x_k)$. In this case the players are said to adopt a *mixed strategy.* Denoting the loss function (of the first player) by Ψ, the average loss function is given by

$$\Psi^P(p_1, p_2) = \sum_{x_1 \in F_1} \sum_{x_2 \in F_2} p_1(x_1) p_2(x_2) \Psi(x_1, x_2)$$

defined on the set $K_1^P \times K_2^P$, where

$$K_k^P = \left\{ p_k : F_k \to [0, 1] : \sum_{x_k \in F_k} p_k(x_k) = 1 \right\}.$$

Theorem *Any duel with a finite numbers of strategy profiles admits a Nash equilibrium made of mixed strategies.*

PROOF Just observe that K_k^P and Ψ^P fulfill the hypotheses of Theorem 2.24. $\diamond$

Notes on Chapter 2

The implicit function Theorem 2.4 is due to Dini [*Opere vol.II*, Ed. Cremonese, Roma (1954)]. Concerning the Newton method, we refer to the paper of Kantorovich [*Acta Math.* **71**, 63–97 (1939)].

Theorem 2.8 and Theorem 2.8 are refined versions of the "method of the successive approximants", envisaged by Cauchy and Liouville, and developed in the most general form by Picard [*J. de Math.* **6**, 145–210 (1890)]. Theorem 2.11 (for finite-dimensional Banach spaces) is due to Peano [*Math. Ann.* **37**, 182–228 (1890)]. Dieudonné's Example 2.12 is in *Acta Sci. Math. Szeged* **12**, 38–40 (1950). Theorem 2.13 is due to Godunov [*Functional Anal. Appl.* **9**, 53–55 (1975)]. The same result for nonreflexive Banach spaces has been obtained by Cellina in a previous paper [*Bull. Amer. Math. Soc.* **78**, 1069–1072 (1972)].

Lomonosov's Theorem 2.16 has appeared in 1973 [the English translation is in *Functional Anal. Appl.* **7**, 213-214 (1974)]. For some time it was not clear whether there could exist operators to which the theorem does not apply. An example in that direction has been found by Hadvin, Nordgren, Radjavi and Rosenthal [*J. Funct. Anal.* **38**, 410–415 (1980)]. However, the problem of invariant subspaces in Banach spaces has been solved (negatively) by Enflo some years later [*Acta Math.* **158**, 213–313 (1987)]. A good reference for the subject is the book of Beauzamy *Introduction to operator theory and invariant subspaces*, North-Holland, Amsterdam (1988).

More details on measure preserving maps and their link with ergodic theory can be found in Walter's book *An introduction to ergodic theory*, Springer-Verlag, New York (1982).

Example 2.17 is due to Banach [*Opérations linéaires*, Monografje Matematyczne **I**, Warszawa (1932)]. Day's Theorem 2.18 is in *Illinois J. Math.* **1**, 509–544 (1957).

A complete reference for Haar measures is Nachbin's book *The Haar integral*, Van Nostrand, Princeton (1965).

The celebrated Theorem 2.23 on non-cooperative games is due to Nash [*Ann. of Math.* **54**, 286–295 (1951)], who obtained for the result the Nobel Prize in Economics. The first striking result in game theory is von Neumann's Theorem 2.24 [*Ergebnisse eines Math. Colloq.* **8**, 73–83 (1937)].

Bibliography

[1] H. Brezis, *Analyse fonctionnelle - Théorie et applications*, Masson, Paris (1983)

[2] K. Deimling, *Nonlinear functional analysis*, Springer-Verlag, Berlin (1980)

[3] N. Dunford, J.T. Schwartz, *Linear operators*, Interscience Publishers, New York (pt.I 1958, pt.II 1963)

[4] R.E. Edwards, *Functional analysis*, Holt, Rinehart and Winston, Inc., New York (1965)

[5] L.C. Evans, *Partial differential equations*, Amer. Math. Soc., Providence (1998)

[6] V.I. Istrăţescu, *Fixed point theory*, D. Reidel Publishing Co., Dordrecht (1981)

[7] J.L. Kelley, *General Topology*, Van Nostrand Co., Princeton (1955)

[8] W.S. Massey, *A basic course in algebraic topology*, Springer-Verlag, New York (1991)

[9] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New York (1983)

[10] F. Riesz, B. Sz-Nagy, *Functional analysis*, Frederick Ungar Publishing Co., New York (1955)

[11] H.L. Royden, *Real analysis, 3rd ed.*, Macmillan Publishing Company, New York (1988)

[12] W. Rudin, *Principles of mathematical analysis, 3rd ed.*, McGraw-Hill Book Company, New York (1976)

[13] W. Rudin, *Real and complex analysis, 3rd ed.*, McGraw-Hill Book Company, New York (1986)

[14] W. Rudin, *Functional analysis*, McGraw-Hill Book Company, New York (1973)